International Edition

The Theory and Practice of Sustainable Engineering

Braden R. Allenby

PEARSON

The Theory and Practice of Sustainable Engineering

Braden R. Allenby
Arizona State University

International Edition contributions by
Sudhir Chella Rajan
Indian Institute of Technology Madras

Prentice Hall
Upper Saddle River Boston Columbus San Francisco New York Indianapolis
London Toronto Sydney Singapore Tokyo Montreal Dubai Madrid
Hong Kong Mexico City Munich Paris Amsterdam Cape Town

Vice President and Editorial Director, ECS: Marcia J. Horton
Executive Editor: Holly Stark
Editorial Assistant: Keri Rand
Marketing Manager: Tim Galligan
Production Manager: Pat Brown
Publisher, International Edition: Angshuman Chakraborty
Acquisitions Editor, International Edition: Arunabha Deb
Publishing Assistant, International Edition: Shokhi Shah
Art Director: Jayne Conte
Cover Designer: Jodi Notowitz, Wicked Design
Full-Service Project Management: Sandeep Rawat/Aptara®, Inc.
Printer/Binder: Courier Westford
Cover Printer: Courier Westford

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsoninternationaleditions.com
© Pearson Education Limited 2012

The right of Braden R. Allenby to be identified as author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

Authorized adaptation from the United States edition, entitled The Theory and Practice of Sustainable Engineering, 1st edition, *ISBN 978-0-13-212799-8 by Braden R. Allenby published by Pearson Education © 2012.*

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a licence permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

Microsoft® and Windows® are registered trademarks of the Microsoft Corporation in the U.S.A. and other countries. Screen shots and icons reprinted with permission from the Microsoft Corporation. This book is not sponsored or endorsed by or affiliated with the Microsoft Corporation.

ISBN 10: 0-273-75216-2
ISBN 13: 978-0-273-75216-5

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library
10 9 8 7 6 5 4 3 2 1
14 13 12 11 10

Typeset in Minion-Regular by Sandeep Rawat/Aptara®, Inc.
Printed and bound by Courier Westford in The United States of America

The publisher's policy is to use paper manufactured from sustainable forests.

To my family, and especially to Carolyn, my patient partner

He, only, merits freedom and existence
Who wins them every day anew.

—Goethe, Faust, lines 11, 575–76

CONTENTS

List of Text Boxes 8

Preface and Operating Manual 9

Chapter 1 SO WHAT IS DIFFERENT NOW? OR, WHY WE NEED SUSTAINABLE ENGINEERING 15

1.1 Introduction 15

1.2 Welcome to the Anthropocene 18

1.3 Evolution of the Anthropocene 21

1.4 The Fingerprint of the Anthropocene: Energy and Water 26

1.5 The Fingerprint of the Anthropocene: Human Imprints on Sinks and Material Flows 35

1.6 Conclusion 45

Exercises 48 • Annotated Bibliography 49

Chapter 2 THEMES OF THE ANTHROPOCENE 50

2.1 Introduction 50

2.2 Globalization and Multiculturalism, Then and Now 51

2.2.1 Multiculturalism and Technological Evolution 57

2.2.2 Changes in Governance 59

2.3 Information and Communications Technology and Systems 62

2.4 Technology and Sustainability 67

2.5 Demographic Trends 69

2.5.1 Urbanization 70

2.5.2 Urbanization and Sustainable Engineering 72

2.5.3 Age-Related Demographic Issues 78

2.5.4 Demographics and Risk 85

Exercises 87 • Annotated Bibliography 88

Chapter 3 COMPLEXITY 90

3.1 Complexity 91

3.2 Simple versus Complex Systems 92

3.3 Different Forms of Complexity 101

3.4 Understanding Complexity 106
3.5 Models, Ideology, and Complexity 107
Exercises 111 • Annotated Bibliography 112

Chapter 4 SUSTAINABILITY 113
4.1 Introduction 113
4.2 Sustainable Development 114
4.2.1 Pillars of Sustainable Development 115
4.2.2 Sustainability as Myth 118
4.3 Sustainability and Resources 122
4.4 From Sustainability to Sustainable Engineering 128
4.5 Sustainability and Global States 132
Exercises 135 • Annotated Bibliography 136

Chapter 5 HOMO FABER: HUMAN HISTORY AND TECHNOLOGY 138
5.1 Introduction 138
5.2 The Railroad as Technology System 147
5.3 Technology Clusters 156
Exercises 160 • Annotated Bibliography 161

Chapter 6 CHARACTERISTICS OF TECHNOLOGY 163
6.1 Introduction 163
6.2 Product Design 163
6.3 Behavior of Technological Systems 168
6.4 The Three Levels of Technology Systems 178
6.5 Concluding Observations 184
Exercises 186 • Annotated Bibliography 187

Chapter 7 INDUSTRIAL ECOLOGY 189
7.1 Introduction 189
7.2 Industrial Ecology 190
7.3 Life Cycle Assessment 198
7.4 Streamlined LCA 205
7.5 Systems Engineering 210
Exercises 213 • Annotated Bibliography 213

Chapter 8 THE FIVE HORSEMEN: EMERGING TECHNOLOGIES 215

8.1 Introduction 215

8.2 The Five Horsemen 219

8.3 The Human as Design Space 239

Exercises 243 • Annotated Bibliography 244

Chapter 9 GREEN CHEMISTRY 245

9.1 Introduction 245

9.2 The CFC Case Study 253

9.3 Antimicrobials and Pharmaceuticals as Earth Systems 256

9.4 Policy Implications 261

Exercises 263 • Annotated Bibliography 264

Chapter 10 SUSTAINABLE ENGINEERING: INFORMATION AND COMMUNICATION TECHNOLOGY 265

10.1 Introduction 265

10.2 Framing ICT 267

10.3 Infrastructure Operations 272

10.4 Services 277

10.5 The Virtualization of Work, Digital Nomads, and the Triple Bottom Line 280

10.6 Virtualization of Work and the Evolution of Institutional Complexity 288

10.7 Conclusion 296

Exercises 297 • Annotated Bibliography 297

Chapter 11 THE FIVE HORSEMEN, MILITARY OPERATIONS, AND NATIONAL SECURITY 299

11.1 Introduction 299

11.2 Thinking about National Security 302

Case 1: Cyborg Insects and Other Surveillance Devices 313

Case 2: Telepathic Helmets 316

Case 3: Lethal Autonomous Robots 319

11.3 Concluding Thoughts 323

Exercises 324 • Annotated Bibliography 325

Chapter 12 THE MACROETHICS OF SUSTAINABLE ENGINEERING 326

12.1 Introduction 326

12.2 Framing Ethics 328

12.3 The Challenge of Macroethics 335

Exercises 340 • Annotated Bibliography 341

Chapter 13 THE ARAL SEA, THE EVERGLADES, AND ADAPTIVE MANAGEMENT 342

13.1 Introduction: Adaptive Management 342

13.2 The Aral Sea 345

13.3 The Florida Everglades 350

13.4 Themes 355

13.5 Adaptive Management Principles 357

Exercises 359 • Annotated Bibliography 359

Chapter 14 EARTH SYSTEMS ENGINEERING AND MANAGEMENT: SUSTAINABLE ENGINEERING AT A PLANETARY SCALE 361

14.1 Introduction 361

14.2 Geoengineering 363

14.3 Urban Design and Management, and High Modernism 367

14.4 Theoretical ESEM Principles 370

14.5 ESEM Governance Principles 375

14.6 ESEM Design and Management Principles 381

14.7 Sustainable Engineering, ESEM, and the Final Principle 387

Exercises 388 • Annotated Bibliography 389

Chapter 15 THE ENGINEER AS LEADER 391

15.1 Introduction 391

15.2 Attitude and Preparation 393

15.3 Lead by Following 397

15.4 Personal Characteristics 398

Exercises 400

Appendix A: Introduction to Sustainable Engineering Generic Syllabus 401

Appendix B: Glossary of Key Terms 413

Index 423

LIST OF TEXT BOXES

BOX 1.1 What Is Sustainable Engineering? 16
BOX 1.2 Exponentials 19
BOX 1.3 Case Study: Corn-Based Ethanol 42
BOX 1.4 Writing for Your Audience 46

BOX 2.1 Monsanto, Genetically Modified Organisms, and Corporate Responsibility 63
BOX 2.2 Lessons of the Autonomic City for Sustainable Engineering 76

BOX 3.1 Plug-in Automobiles and Biofuel: Why Not to Fall in Love with a Technology 104

BOX 4.1 The President's Council on Sustainable Development 119
BOX 4.2 Reserves and Resources 125
BOX 4.3 Toxicity, Hazard, and Risk 133

BOX 5.1 Aspects of Technology 139
BOX 5.2 Technological Determinism 143
BOX 5.3 The Technological Trajectory of Money 145
BOX 5.4 Perception and Technology 159

BOX 6.1 Levels I and II, Revisited: The Shop Floor 182

BOX 8.1 Transhumanism 217
BOX 8.2 Who Defines the Playing Field? Conservation Biology and the Biodiversity Crisis 224
BOX 8.3 Integrating the Brain with the Machine 228
BOX 8.4 Using Metaphors and Analogies 234

BOX 9.1 Principles of Green Chemistry 248
BOX 9.2 Arsenic Management in the United States 251

BOX 10.1 A Telecom Firm Perspective on Sustainable Engineering 268
BOX 10.2 Design for Environment 273
BOX 10.3 Scenarios and Scenario Planning 281
BOX 10.4 ICT at the Bottom of the Pyramid 293

BOX 11.1 Hard and Soft Power in the Modern World 304

BOX 12.1 Respect and Reciprocity: The Golden Rule 329

BOX 14.1 Mining as ESEM 385

PREFACE AND OPERATING MANUAL

When in the Course of human events it becomes necessary for one people to dissolve the political bands which have connected them with another and to assume among the powers of the earth, the separate and equal station to which the Laws of Nature and of Nature's God entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the separation.

—American Declaration of Independence

No one confuses a textbook with the American Declaration of Independence. Nonetheless, one may take guidance from that document: When the world and "human events" change to the extent that old ways of framing problems and designing solutions become increasingly inadequate, and a different kind of book is necessary, readers are owed an explanation. In this case, one has a book purporting to be primarily an engineering textbook, yet even a cursory perusal reveals that there are more words than numbers; more sentences than equations; more qualitative than quantitative exercises; more policy observations than algorithms. What sort of engineering text is this, then, and what possible explanation for its obstreperous deviancy?

To begin with the second question, which will segue into the first: There are some fundamental issues with engineering, engineering education, and sustainability as an objective policy and design criterion—not to mention changing demands on the engineering profession, technology as a practice and a social force, and our society—that demand corresponding change in education.

1. Engineering education is increasingly broken. It tries to teach ever more complicated technologies, and problem solving in ever more complex social and environmental contexts, in a four-year, highly reductionist, and quantitative program that has been inadequate for years.[1] With the

[1]You needn't take my word for it. Take a look at the recently released 2010 book, *Holistic Engineering Education: Beyond Technology,* edited by Domenico Grasso and Melody Brown Burkins (New York: Springer). It has chapters by people such as Jud King, who used to be provost and senior vice president for Academic Affairs for the University of California; Jim Duderstadt, who used to be dean of engineering and then president at the University of Michigan; and a number of other industry and engineering luminaries who, in one way or another, in elegant and direct prose, clearly highlight the inadequacy of current engineering education in the United States, Europe, and elsewhere, and its inability to produce graduates adequate to the ever increasing burden of modern engineering practice.

accelerating evolution of technology across virtually its entire frontier—biotechnology, nanotechnology, information and communication technology, robotics, and applied cognitive science, to name some of the most salient—the education we are used to providing no longer produces the engineers that society needs.

2. The sustainability discourse is becoming increasingly ever more damaged. Serious and legitimate concerns about deep systems perturbations—climate change, shifts in biodiversity, urbanization and aging demographics, energy systems, economic behavior—are trivialized by a dialog that seems to increasingly turn in on itself, and becomes ever more a question of faith and belief among the self-anointed (despite being reified as "sustainability science" and for that matter "sustainable engineering"). Society is ill-served by the capture of sustainability issues by dogmatic and ideological factions.
3. Institutional, governmental, and social ability to deal with the complexity of the **anthropogenic Earth**[2]—the world of human activities, desires, and life that we have already built around ourselves, a world where systems dynamics are increasingly determined by human activities, intentional and unintentional—is broken. With issue after issue, in country after country, an inability to deal with complexity is tending to lead to shrill ideology and fragility as culture wars, within countries and among human belief systems, becomes the norm. Especially as technological change accelerates, destabilizing existing beliefs, cultural norms, and social morays, the reaction to complexity is not adaptation, but denial and retreat into simplistic frameworks—fundamentalism of all kinds. If you want to solve problems in this complex world, however, you don't have that luxury, whether you're an engineer, a lawyer, a business executive, or an urban planner. But neither, because of the way we currently educate you, do you have the tools you need.

This book alone will not fix these and related issues, of course—even the usual authorial hubris can't pretend that. It tries, however, to do several things. Most important, it tries to sketch out the cultural, social, institutional, and environmental context within which engineering and, more broadly, technology systems, are now situated. This obviously will not turn lawyers into biotechnologists, or electrical engineers into sociologists, and it would be foolish to try to do so. One hopes, however, that it will begin to provide some frameworks within which understanding, communication, and solutions to highly complex problems with significant technological dimensions, can occur. It also tries to break down some of the barriers that disciplinary boundaries and

[2]Throughout the text, certain key terms will be marked in bold type. For the convenience of the reader, those terms are collected and briefly defined in the glossary provided in Appendix B.

reductionist methodologies, have created, not in the name of replacing reductionism, which remains a critical tool for scientists, engineers, and society, but with the goal of generating more capable professionals. Such individuals should be competent in their chosen field of course, but also able to integrate across other disciplines to address complex adaptive systems.

As such, this text is not intended just for engineering professionals. To be sure, despite its qualitative appearance, this is an engineering textbook, but it is also a textbook for any student, of any age, who wants to better understand the strange new anthropogenic world we have created for ourselves and those who share the planet with us—and who want to rationally, ethically, and responsibly respond to the challenges and emergent behaviors that world is now tossing up for us. After all, it is not just the engineers designing the products and services that we use who need to understand more about technology, context, and systems implications: it is also those who write and pass the laws and regulations that provide us guidance and embody our ethics; who create policies for cities, states, and countries; who educate all sorts of students to do all the wonderful things students do when they leave their colleges and universities; who run the firms, research parks, and venture capital firms that ride new technologies to market and into our lives at ever faster rates; and who by their consumption and activity decisions shape the globe. "Sustainable engineering" is in the title of the book, but it requires more than just engineers to make that desirable capability part of our human toolbox: It requires all of us. So although this is a book about sustainable engineers, it is meant for a broader audience. We all live in that ever more human world that we have built together over many centuries and through many cultures, and it cannot be designed and built into the future as we wish it just by engineers. In short, this is not just an engineering text for engineering students, but also an engineering and technology text for nonengineers.

A book such as this also necessarily delves into issues and debates that are highly charged. Accordingly, I also note in advance that, while I have tried to be neutral in some very conflicted areas, that effort might not always be successful. In my case, I come from an industry background, and have also spend time in government at a national defense laboratory and in academic positions, a set of experiences that no doubt provides some background for, and shaping of, my perspectives. Moreover, if there is to be coherence in discussing very complex subjects, one must adopt particular perspectives: cubism, with its multiple superimposed perspectives, works in art, but it is highly questionable when applied to engineering texts. In those instances, then, where description lapses into assertion, whether you as reader agree with me or not, you should always remember and respect the validity of other perspectives. Authors cannot escape their experiences: if they help you understand more deeply even as you disagree with them, they are worthwhile; if they are simply biased, learn what you can from them, reject the bias, and move on.

Even a cursory glance at the table of contents may raise other questions. Why, for example, is there a chapter on military technology? That's not part of the typical text, for sure. Nevertheless, the disciplinary boundaries and categories that we take for granted are shifting these days, more rapidly and unpredictably than they have in a long time. Implicit assumptions about what matters and what doesn't need to be rethought. Consider military operations and technology in that light. We live in a world where, for many reasons, military necessity and pressure drives technological evolution, and if emerging technologies are important, then we need to be alert to developments in the military arena, because that's where a lot of these technologies will be coming from. Moreover, the strong boundary line dividing military and civilian technologies was never absolute, and it's growing weaker. Not only do military systems increasingly include civilian components and assemblies, but changes in military doctrine in some countries and the rise of terrorism targeting civilian environments mean that civilian technologies and systems can no longer be thought of as isolated from military challenge. If a terrorist wanted to attack American information and communication networks, for example, why should he try to attack the Pentagon, which is relatively well defended in cyberspace, when he could much more easily attack the information systems that are increasingly critical for the function of U.S. transportation, energy, water, financial, and other systems—all designed by engineers with little if any knowledge of cybersecurity? More subtly, it used to be that many students had at least some familiarity with military culture, technology, and strategy because of mandatory service requirements in many countries. Now, the rise of volunteer militaries means that we can no longer rely on that almost universal experience; a connection that used to be made through life experience now needs to be made through education. Both as a symbol of shifting categories and as a substantive topic for all educated technologists, this chapter on military engineering is a necessary one. It may be novel in such a textbook, but the world has changed. And so must education.

A more basic question, perhaps, is how students facing an already overloaded course load can jam a course on sustainable engineering and the implications of technology systems on top of everything else they must study. This is a serious challenge, especially for the highly structured engineering curriculum, but it is manageable.

There are two ways that this subject matter and this text might be taught effectively. The first, which has a long tradition in engineering education at least, is to use this book to provide modules in existing courses, or as a supplementary text in sustainable engineering, green engineering, industrial ecology, sustainability law and policy, and environmental courses. This is worthwhile, but of course is fairly incremental.

The second is as the basis for a course in itself, which is what the book has been designed for. This is not as outrageous as it seems: Almost all courses of study, including engineering, require distributional credits outside of the primary disciplinary area. In many cases, such courses are simply the easiest that the student can find outside the major (e.g., "The Feline Form in Comic History"). Wouldn't it be better to offer a course that meets distributional requirements, but does so by augmenting the understanding of the student—whether in engineering, policy, environmental studies—with a working knowledge of technology, technological change, and the associated dynamics of social, cultural, and environmental systems? Then the class wouldn't need to fit within the core engineering tracks, but could substitute as a more useful distributional course for others that may be somewhat idiosyncratic and ad hoc. Moreover, many schools are realizing that just as their engineering programs need a sustainable engineering component, their liberal arts, business, and policy students need some exposure to technology and technology systems, and the engineering environment from which they flow. After all, given the ever more important role played by technology in our world, not just technocrats, but every citizen, needs some familiarity with where and how technology comes to be.

Just as this text is nontraditional, so too is a course structured around it. Simply saying such a course might be useful is accordingly inadequate. For that reason, a suggested course syllabus is included in Appendix A. Obviously, there are many ways such a course can be structured, and the intent is not to suggest that this is the best way to do so. (To begin with, not all schools will have the same schedule as laid out in the Appendix.) Rather, the intent is to offer a complete course instead of just a textbook, thereby easing adoption of what to many people may be unfamiliar, at least in an engineering curriculum. The other reason the syllabus is included as an appendix rather than as a part of the instructor's guide is because it may be helpful for readers of this volume, including students, to understand how the material translates into a course format. In a sense, this is metalearning: In addition to learning the substance of technical sustainable engineering, it is useful for the sustainable engineer to understand sustainable engineering itself as a process of change, and learn from that as well. Sustainable engineers will often be change agents, so learning how change may be introduced into an institution, be it a firm or a discipline, is part of the remit.

A word on style. Because this is intended to be used as a textbook, it is not heavily footnoted. That should not hide the obvious fact that any broad discussion such as this necessarily draws on the accumulated wisdom of many, many people, the author probably least among them. I have attempted to indicate the major sources that have helped me in various areas by adding an annotated bibliography to each chapter, where books that I have found useful

and informative are identified. Where I have quoted directly or relied particularly heavily on an individual source, however, it has been noted either in the text or in a footnote.

In this regard, the outstanding work of Carolyn Mattick, who was instrumental in researching the graphics in this book and in helping me create the course based on it, cannot be overstated. She has displayed a passion for understanding the data and the issues, and for helping to create new and innovative curricular materials, augmented by a working ethic that is all too rare these days. I deeply appreciate her contributions to this book and our ongoing collaborative work on sustainable engineering; she cannot be blamed for the outcome, but without her it would be a far less interesting and complete effort. If it had happened at all. My thanks.

My thanks also to those many, many people who I have worked with and learned from, and who have challenged me. Working across boundaries is difficult in part because one never knows new intellectual terrain as well as those who already inhabit it, and I have been very fortunate to have found colleagues who have been gracious enough to show me the many errors of my ways constructively and in good spirits. I also thank those from whom I learned via books, articles, and lectures: It may not have been personal, but their congealed knowledge, transmitted in various ways, has been and continues to be not just the basis for this work, but a source of joy and inspiration. I won't mention names, because I would inevitably leave out so many that I would be overwhelmed with guilt for the rest of my days; moreover, it would be unfair to those who helped me with such fair spirit to associate them with what I turned their wisdom into. For that, I must take full responsibility.

This book benefited from the work and conferences conducted by the Center for Sustainable Engineering, supported in part by the National Science Foundation through Grant DUE-0442618. Additional support was provided by the Lincoln Center for Applied Ethics at Arizona State University. Of course, the opinions, findings, conclusions, and recommendations expressed in this material are mine, and do not necessarily reflect the views of either of these fine organizations, or of the United States government.

Brad Allenby

www.pearsoninternationaleditions.com/allenby

CHAPTER

1 So What Is Different Now? Or, Why We Need Sustainable Engineering

It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is the most adaptable to change.

—Charles Darwin

1.1 INTRODUCTION

Every once in a while, engineering undergoes profound shifts. In the late 1700s and early 1800s, for example, civil engineering as a discipline was differentiated from military engineering on the one hand, and architecture and craft practices, such as masonry, on the other. Among the reasons for this shift were the professionalization of technology and scientific activities that accompanied the Industrial Revolution, as well as an increasingly apparent difference between military activities, which had been the traditional focus of engineering, and application of those skills in nonmilitary, "civil," engineering projects. Electrical and electronic engineering, on the other hand, became a recognized field only in the late nineteenth century, for the obvious reason that the technologies with which such engineers worked—notably the telegraph and nascent electric devices, including generation and transmission facilities—didn't exist until then. Similarly,

aeronautical and aerospace engineering couldn't be a recognized field of practice until the technologies that enabled flight and access to space developed.

So why sustainable engineering, and why now? This turns out to be a surprisingly difficult question, which is perhaps approached by asking a different one: What is different about this world now that might call forth a different kind of engineering? And here we find much to think about. (Meanwhile, Box 1.1 provides a short and high level answer to what sustainable engineering is.)

BOX 1.1 What Is Sustainable Engineering?

There are no easy answers, and the observant reader will notice that this entire book is, in a way, a long essay on this complicated question. But we should begin with at least a grasp of the essentials.

Let's begin with the second term first. "Engineering" is structured problem solving, using the methods and tools developed over decades and centuries by earlier generations of engineers. These tools are usually quantitative rather than qualitative—that is, they tend to require numerical data as input, to involve numerical algorithms, models, and methods, and to generate numerical output. Moreover, engineering is further constrained because engineers have to develop solutions that actually work and that solve problems in the real world. Such solutions may be hardware—for instance, product and component designs, built infrastructure such as buildings, roads, water, and sewage systems, and artifacts of all sorts. Increasingly, however, they are more complex than that, including **wetware** (biological components), **software** (information and programming components), **orgware** (institutional and organizational design components), or other more exotic **x-ware**. In fact, one of the reasons engineers now have to worry about **sustainable engineering** is precisely because the problems that they are expected to solve have grown far beyond the traditional boundaries older generations of engineers were accustomed to. Fundamental changes in the problems that engineers are asked to address thus change engineering itself.

Sustainable is a more difficult term, primarily because there are no universally accepted definitions. Although we deal with "sustainability" and "sustainable development" in more detail in the text, it is important to emphasize several aspects of sustainability at the very beginning. Most important, in almost all formulations sustainability includes two important themes: (1) appropriate concern for natural systems ("the environment"), and (2) appropriate concern for social equity ("equity" in this case means *fairness*, a term that is itself subject to many different interpretations).

For our purposes, this inclusion of dual themes immediately lays out several critical challenges. First, because of many changes in the world—more people, doing more things, affecting more natural systems, using more resources, creating far more complicated technologies, with effects emerging at constantly larger temporal and physical scales—solving problems has become more complicated in the past few years. In particular, what *sustainable engineering* implies is that the sort of problem solving that engineers are being asked to do now includes far more consideration of natural and environmental systems, and of social systems and institutions, than it did previously. This unavoidably changes what engineers need to learn to do. Fear

not. It does not mean that you also need to be an environmental scientist and a sociologist; however, it does mean that for you to be an effective problem solver, you need to be familiar with some of the trends and issues with which these and other experts in their particular fields deal, be able to at least find common ground, and speak a common language when you are placed on a design team with them.

Second, this accelerating complexity and emphasis on sustainability puts a particular burden on engineers. As an activist, for example, you don't need to include all viewpoints; in fact, you often won't because, after all, you are focusing only on the ones of interest to you. So an activist's solutions can be simple: "Just do what I want you to do." But such "solutions" seldom work in a real world full of many different worldviews and interests; as an engineer, you don't have that luxury. You have to create solutions for the real world, and in the real world not everyone agrees with every activist. You need to balance environmental concerns, social concerns, economic concerns, technical constraints and considerations; in short, you need to make something that works, not just something that meets a particular ideological demand.

Third, as an engineer you must always learn to think in terms of systems. This is far harder than it sounds, especially as systems expand to include many different technological domains (wetware and software, for example, and the emerging technologies discussed in Chapter 8), and on top of that, new environmental and social domains. But systems engineering has been with us for some time, and with insights gained from that experience, as well as from nascent fields such as industrial ecology and adaptive management, we can begin to ensure that our working view of systems corresponds to the problems we're charged with solving. For example, if I'm asked to design a toaster, I probably don't have to get too wrapped up in complex questions of redesigning electrical grids, or economic and social analysis, much less environmental engineering (design for environment guidelines such as "minimize use of toxics" and "reduce energy consumption in use" are fairly common these days). But if I'm assigned to engineer the Everglades in Florida for long-term viability . . . well, that's a handful, and it goes way beyond simply building "good" drainage canals. In fact, those canals may now be part of the problem.

So why sustainable engineering? Because the world has changed, because it is continuing to change more and more rapidly, and because being an engineer in such a world requires different skills and knowledge than it once did. But make no mistake: A good engineer is one who is able to best solve the problem presented in his or her age, with the culture, constraints, materials, and technologies of that time and place. In this very fundamental way, sustainable engineering is simply good engineering in an ever more complex world. So, although in one way sustainable engineering is new, it is in another way as old as the profession itself; it is engineering reconstructing itself for the age of the human Earth.

That last point also explains why sustainable engineering—good engineering—is increasingly important. Our world is increasingly dominated by human activity and by our technology systems—thus, engineering matters even more now than before. This means that sustainable engineering is not just your professional responsibility and an important component of your intellectual capital, it is also a critical competency for society and for our future.

1.2 WELCOME TO THE ANTHROPOCENE

We live in a world that is fundamentally different from anything that we have known in the past. It is a world increasingly dominated by the activities, technologies, and cultures of one species—the human species. So welcome to the first terraformed planet: not Mars, or a moon of Saturn, but Earth. In fact, many scientists are now calling this latest geological period the **Anthropocene**, or, roughly translated, the Age of the Human.

This sounds extreme at first, but a few examples will help illustrate the basic idea. If you were an alien exploring our corner of the galaxy and saw Earth for the first time, for instance, what would you notice? For one thing, you'd see large parts of the land area that are obviously managed, and networks of physical structures—roads, canals, power grids, air routes, the Great Wall of China—that completely covered the globe. Using your infometers, you would immediately detect vast quantities of information shifting around the world, managed by a single species for its benefit; perhaps your instruments would be affected by the torrent of broadcast information emanating from the planet (sophisticated alien speedometers being thrown off by reruns of television shows is an interesting thought). As the world shifted into night, you would see vast areas of artificial light. You might detect that the dynamics of the atmosphere and ocean circulation patterns are increasingly structured by human activity; you would certainly see the contrails of aircraft and lingering smoke trails of ships, crisscrossing Earth. You would note that an increasingly large percentage of photosynthetic product was appropriated by humans: in some areas for eating (agriculture and pasture), in some areas for aesthetics (e.g., lawns in rich countries), in some areas as a design choice (e.g., keeping the Everglades "wild" as a design decision), and in some areas using designed life forms (**genetically modified organisms**, or "GMOs"). In short, whether it was intended or not, you would see an anthropogenic—"human derived"—Earth. And the key to this Earth is integrated systems of human societies and cultures; human design, especially built infrastructures and products; and nature.

Here is the short answer, then, to why engineering is changing yet again, and why the focus is on sustainable engineering: It is now a human Earth. And if we're going to be ethical, responsible, and rational in how we live on it, continue to change it, and build our futures, we need to evolve our engineering practices beyond simply better widgets. Engineers need to become far more proficient at understanding the social, cultural, environmental, and technological systems within which they practice their profession; more important, they need to be ready to assume intellectual and policy leadership in a world significantly affected by the technologies they create. Sustainable engineering, therefore, is nothing more, but nothing less, than learning how to engineer responsibly and professionally in the Anthropocene. But to do that, perhaps some idea of why and how we got here is appropriate.

In doing so, the reader will note that many graphs in this chapter and in the ones that follow, such as the ones illustrating human population growth or the increase in computing power over time, seem to curve dramatically upward at some point. Moreover, many systems, from population growth of a species introduced into a new environment to technologies over their lifecycle (Chapter 6), assume an "S" shape: that is, they start slowly, accelerate rapidly, but as they approach their peak begin to level off, forming the shape of an "S." Exponential growth, limited eventually by environmental resources or changes in economic systems, often display such characteristics; Box 1.2 briefly introduces the mathematics and curves common to such systems.

BOX 1.2 Exponentials

There's an old story about a Chinese emperor who is presented with a chessboard, which of course has sixty-four squares on it. The emperor is moved by the gift and asks the young lad who has presented it to him what he wants as a reward. The lad replies that he does not wish to be greedy, and would simply accept rice that is doubled for each square: one grain on the first square, two on the second, and so forth. The emperor is surprised at his lack of ambition, but agrees, and calls for rice to be brought. Everything goes well for a little while, but by the twenty-first square over a million grains is required, and by the forty-first square a billion grains is required. There is not enough rice in the world for the final squares, and the emperor realizes he has been a victim of exponential growth. Very simply stated, an exponential equation is one where a variable appears in an exponent, such as the "x" in the equation given in Figure B1.2.1. When graphed, it gives the

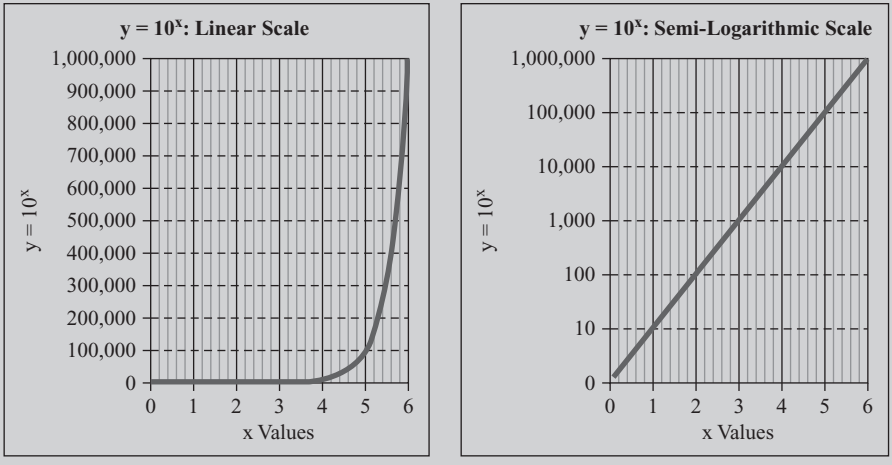


FIGURE B1.2.1 Exponential Function. Both of the above graphs show the same function, $y = 10^x$. However, the graph on the left displays it on a linear scale whereas the graph on the right shows the same information on a semi-logarithmic (base 10) scale.

characteristic upward acceleration noted in the linear scale figure below for positive values of the right-hand side of the equation. When such a curve is plotted on a semi-log scale, it changes into a straight line (which means that when you're looking at any graph in this book, or anywhere else, make sure you know what the values are on the y-axis).

But few systems maintain exponential growth for a long time: The universe is full of constraints. For flour beetles, it's the end of flour; for a product, it's the arrival of a replacement; for a financial market, it's saturation of demand. Eventually, the number of transistors on a chip can't be doubled anymore. Thus, what one often sees is the "logistic curve," or the **S curve**, where the constraint (in the equation in Figure B1.2.2, it would be *K*) acts to slow exponential growth, which eventually comes close to, but never actually reaches, its limit.

The logistics curve, however, suggests that a plateau is reached. This is infrequently the case. After all, if the constraint on flour beetle population growth in a new bin of flour is the flour, the population doesn't just level off: Rather, once the beetles have consumed most of the flour, they'll start dying. In other words, it's a crash. Thus, Figure B1.2.3 suggests a number of possible scenarios once the logistics curve begins to level off. One could have a crash, either back down to baseline or to a lower equilibrium level. One could have a new logistics curve grow from the plateau of the old one. This is a pattern that some have suggested characterizes human population growth, which expands as new core technologies such as the "green revolution" in agriculture create, in essence, a new and higher carrying capacity. Alternatively, one could oscillate around some central value. Isolated populations of predator and prey sometimes display such patterns: a low level of predators leads to an increase in prey, which in turn

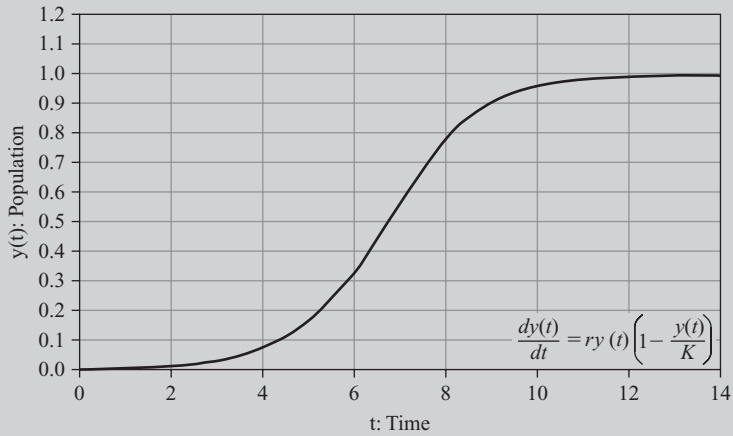


FIGURE B1.2.2 Typical Logistic Growth Curve. In the equation, *r* represents the population growth rate and *K* represents the carrying capacity of the environment.

leads to more predators, which overconsume the prey, thus collapsing their population and, as food runs low, leads to a crash in predator population as well—and the process starts over.

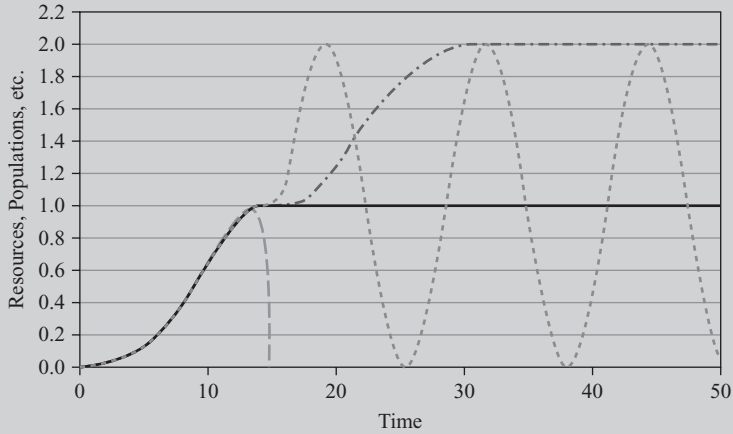


FIGURE B1.2.3 Future Scenarios. This figure illustrates some of the possible behaviors after the "S" curve plateaus. Which one actually occurs depends on the dynamics of the particular system.

1.3 EVOLUTION OF THE ANTHROPOCENE

Although human changes to Earth have been going on for a long time, it was the Industrial Revolution in the 1700s, and associated changes in human demographics, agricultural practices, technology systems, cultures, and economic systems, that really led to a very different Earth. Any human activity, of course, has some effect on the immediate environment—that is true even of early hunter–gatherer society. But there was a step change in these effects when humans gave up hunter–gatherer lifestyles for agriculture, when people began to uniquely affect their local and regional environments: land use, water use, agricultural and sewage runoff, and the like. The second step change in human evolution occurred with the Industrial Revolution: global technologies enabled far more intimate co-evolution of human and natural systems; increasing interconnectedness among human, natural, and built systems meant that new patterns of activity rippled across not just regional, but global systems. With the steam engine, for example, came railroads and steamships, and global consumption patterns that were able to reach into continental interiors to create wheat fields where swamps and forests once reigned.

TABLE 1.1 Human Population Growth and Core Technology State[1]

Age	Population Level (in Millions)	Global Technology State (Core)
1,000,000 BP[2]	0.125	None
10,000 BP	4	Beginning of agricultural revolution
2,000 BP	100	Agricultural
1,000 BP	300	Agricultural
500 BP	450	Enlightenment; Beginning of modern science and technology
100 BP	1,600	Heavy engineering (e.g., railroad)
50 BP	2,500	Mass production and consumption (e.g., automobile)
Today	6,500+	Information/biotechnology society

[1]Notice how rapidly the growth in population accelerates with the Industrial Revolution. Also, always remember that when numerical data on such broad systems are presented, it is best to regard them as approximate, because of inevitable uncertainties in their estimation and aggregation.

[2]Before Present.

Table 1.1, for example, illustrates the basic trends in human population, and makes the fundamental point that the growth in human population, especially with the technological advances that characterized the Industrial Revolution, has been extraordinary. It also makes another important point, however: Human population is a function of, among other things, the technologies humans have available to support themselves; those technologies are, of course, the essence of engineering. From the beginning, then, the importance of understanding technology not just as the widgets that many engineers work with, but as important earth systems with implications for many aspects of human and natural systems, becomes clear.

Of course, it is not just population that matters, but what all those people do, and this can be captured in economic numbers. Figure 1.1 provides an overview of global economic history, and what it shows is not just dramatic growth in world **gross domestic product** (GDP, which is the economic value of the goods and services produced in an economy), but even in GDP growth per individual. What that says is that, even though world population has gone from about 450 million in 1500 to about 6 billion in 2000, economic activity per person has more than kept pace, going up by about an order of magnitude over that time (from 1.0 to 7.57 on a relative basis).

Figure 1.2 presents the history of global economic growth over the last 2,000 years in another light. For centuries, changes in technology, which

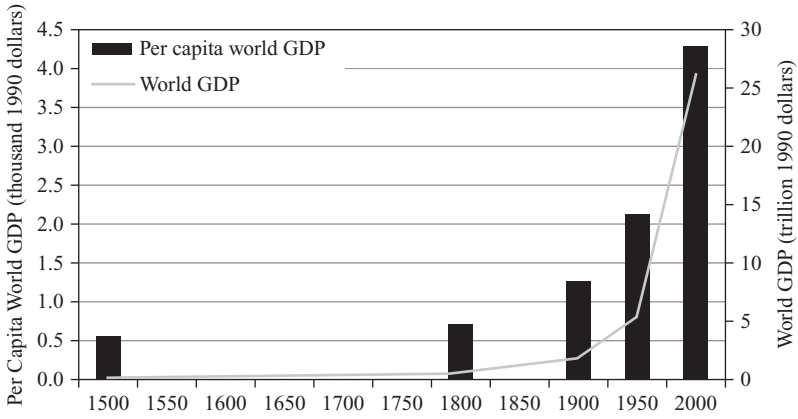


FIGURE 1.1 Global Economic History. Economic activity has gone up dramatically not just on a worldwide basis, but on an individual basis, in spite of strong population growth. Such a rapid growth in economic activity and population, taken together, implies dramatically expanded interactions between human and natural systems; because of the role of technology in economic activity, technology systems are often among the important mechanisms linking population, built environments, natural systems, and, of course, other technology systems.

(Data from: De Long, J. Bradford. "Estimates of World GDP, One Million B.C.—Present," 1998: accessed http://delong.typepad.com/print/20061012_LRWGDP.pdf; McNeill, J. R. *Something New Under the Sun: An Environmental History of the Twentieth-Century World.* New York: W. W. Norton & Company, 2000; United Nations Statistics Division. "GDP [Current US$]." United Nations. http://data.un.org/Data.aspx?q=gdp&d=WDI&f=Indicator_Code:NY.GDP.MKTP.CD: accessed June 4, 2010; United Nations Statistics Division. "Per Capita GDP at Current Prices—US Dollars." United Nations. http://data.un.org/Data.aspx?q=gdp+per+capita&d=SNAAMA&f=grID:101;currID:USD;pcFlag:1: accessed June 4, 2010.)

economists generally measure as increases in productivity, were in rough equilibrium with population levels, leading to a situation where the economic status of most individuals remained roughly equal, generation after generation. Moreover, although there were usually a few more wealthy and powerful individuals, these were so few compared to the multitude that, for all practical purposes, there was a rough and brutal equality between and among generations (a major theme of the sustainability concept; see Chapter 4). New research on this subject suggests that these communities were not the gentle, bucolic, sustainable cultures that modern environmental romanticism has suggested, but in fact were characterized by high levels of disease, violence, and homicide.

The Industrial Revolution changed all that, however. In essence, the rapid takeoff of technology and economic production produced temporarily unlimited resources for the human species, which then began to reproduce in the exponential S curve that is familiar to ecologists. Thus, for example, the growth of population in the United Kingdom, the origin of the Industrial Revolution, looks strikingly like the growth of yeast introduced into a new and nutrient-rich

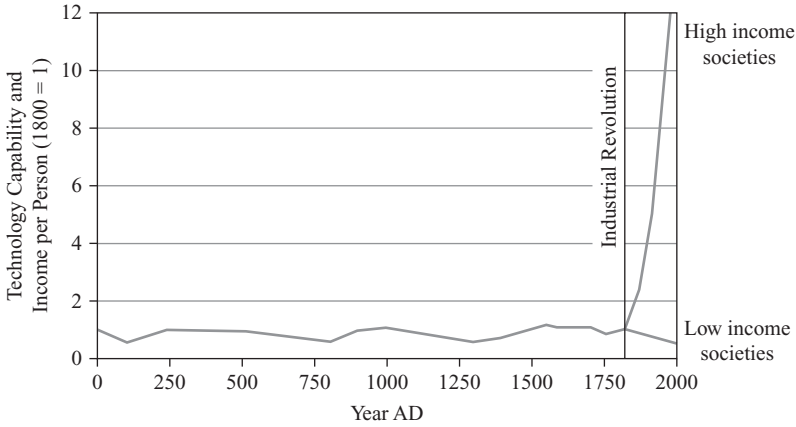


FIGURE 1.2 Global Economic and Technological History Over the Past 2,000 Years. The stable economic performance indicates that technological advances, which were occurring but at a slow pace, were in general balanced by population growth over long periods of human history. It is notable that for most of that time, virtually all humans lived at essentially the same level, a state that changed with the Industrial Revolution and led for the first time to a dramatic divergence between the average income and access to technology, per individual in high-income as compared to low-income societies. This relatively new and substantial gap is sometimes called the "Great Divergence." It is striking that no high-technology society is in the low-income category (except in the aftermath of some catastrophic event such as war, and even then the resilience of high-technology societies tends to be much greater than low-technology societies), and that no society that has not, in some way, learned to use high technology is in the high-income category.

(Data from: Clark, G. *A Farewell to Alms.* Princeton, NJ; © Princeton University Press, 2007. Reprinted with permission.)

environment (compare Figures 1.3 and 1.4). This should not be taken to imply similar dynamics, as the two systems are obviously very different, but it does suggest that one way to understand technology is as a mechanism that creates "new" resources for growth (which is why accelerating technological evolution can create the wealth evident in Figure 1.2 in high-income societies).

Much of this book is about technology, engineering, and social and environmental values, so it is useful to understand clearly and at the very beginning some of the reasons why technology matters. Figures 1.3 and 1.4 thus introduce us to an important and fundamental point: At a very broad scale, technology—taken not just as physical artifacts, but as including the information and social systems of which those artifacts are a part—matters because it is a major determinant of how many humans the Earth can support. Thus, Figure 1.5 presents an overview of human history, associated with the three major global technology states: the Neolithic revolution, the agricultural revolution, and the Industrial Revolution. Moreover, the importance of future patterns of technological evolution is hinted at by the uncertainty of the future path that population might take, which is determined in part by how technology evolves.

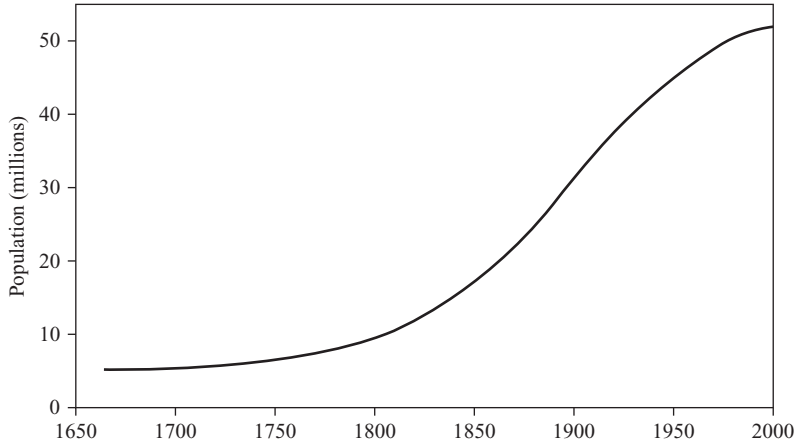


FIGURE 1.3 Population of England and Wales, 1650–2000.

(Data from: Usher, A. P. *A History of Mechanical Inventions*. New York: Dover Publications, 1954; UK Statistics Authority Web site: www.statistics.gov.uk.)

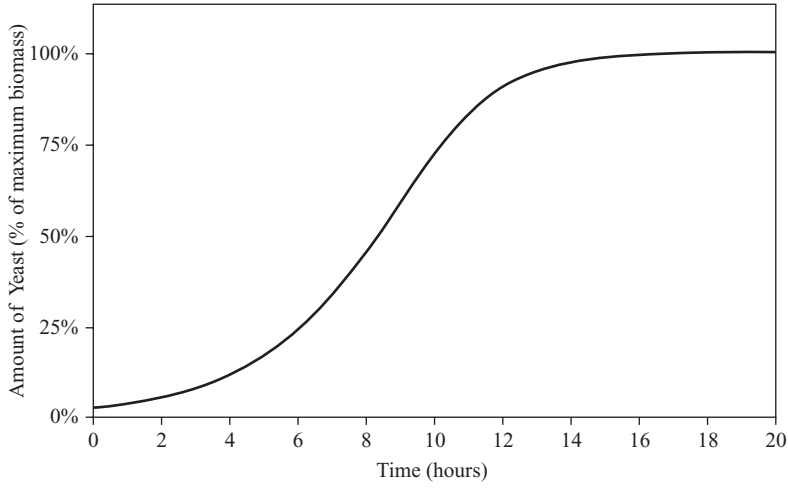


FIGURE 1.4 Typical Growth of Yeast. Figures 1.3 and 1.4 compare the growth of yeast in a new culture, where nutrients are essentially unlimited, with population growth in the United Kingdom over the period of the Industrial Revolution. The two curves arise from similar dynamics: technological change and economic growth as a result of the Industrial Revolution basically created sufficient "new" resources to enable dramatic human population growth in the United Kingdom. Of course, the human pattern integrates very complex and significant underlying cultural, institutional, economic, and demographic changes as well—for example, from a predominantly rural, agricultural economy to a predominantly urban, industrialized economy, and the growth of the United Kingdom as the world's major trading nation. But at a macro level, the similarity of the two growth patterns is suggestive and striking.

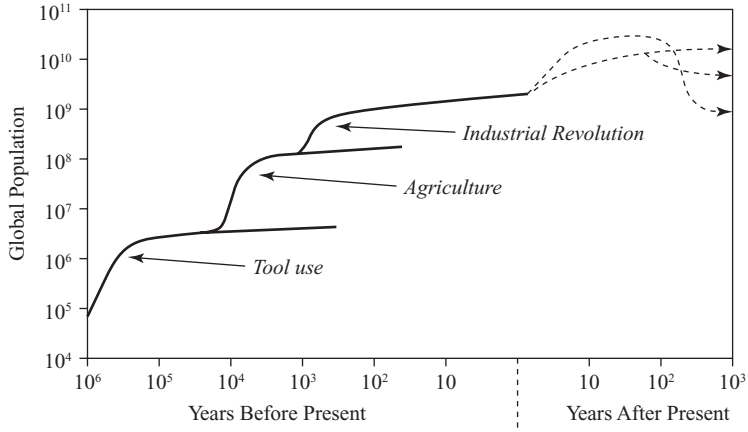


FIGURE 1.5 Global Population. This graph illustrates the relationship between global technological competency and major technology systems: Paleolithic, the agricultural revolution, and the Industrial Revolution. The numbers behind such graphs are, of course, highly approximate, and although the basic relationship is fairly strong, the curves themselves should not be regarded as definitive.

(Based on: Graedel, T. E. and B. R. Allenby. *Industrial Ecology.* 2nd ed. Upper Saddle River, NJ: Prentice Hall. © 2003, p. 4. Reprinted by permission of Pearson Education, Inc.)

It is always important to bear in mind that economic growth is different from the distribution of wealth and assets. Thus, both the distribution of technology and related economic benefits have not been uniform either within or across countries; indeed, for the first time in human history a massive gap opened up between the countries that developed as a result of the Industrial Revolution and those that didn't. This gap is so prominent, and gives rise to so much speculation and concern, that it has been labeled the **Great Divergence** in the literature (this divergence is quite evident in Figure 1.2). People who are well off today are more well off than ever before in history; but the poorest people today are as poor as any of those known to history.

1.4 THE FINGERPRINT OF THE ANTHROPOCENE: ENERGY AND WATER

The economic and population growth, as well as the technological change, that we call the Industrial Revolution had a number of obvious and significant impacts on natural systems. Perhaps the most apparent are the implications for energy flows and water, both critical categories for technological and economic innovation. Take the change in energy consumption over this period, as shown in Figure 1.6. Energy is pivotal for two reasons. First, it is a critical enabler of material transformation and thus economic growth. Second, the burst of energy relied on by humans to support the Industrial Revolution and consequent explosion of human economic and population growth around the globe has relied primarily

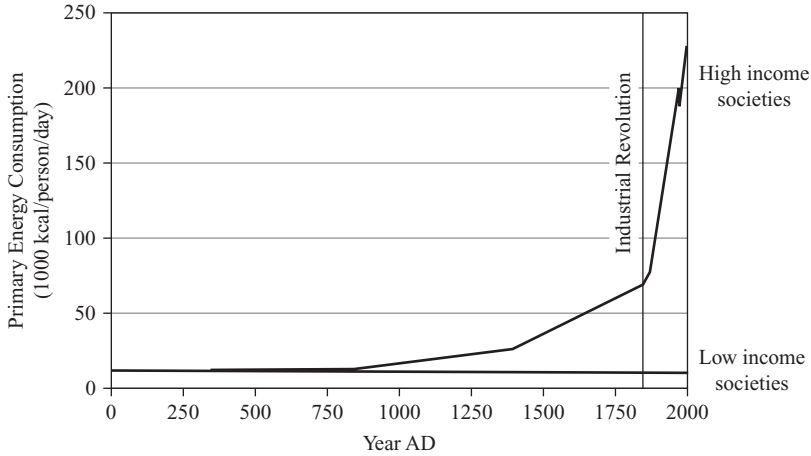


FIGURE 1.6 Global Energy Consumption History Over the Past 2000 Years. Note how the shape of the curve mirrors that in Figure 1.2, illustrating the important relationship between economic development and availability of adequate, inexpensive energy resources. Similarly, note the existence of "The Great Energy Divergence," between those societies with substantial energy consumption per capita, such as the United States, and those with energy consumption per capita similar to that which has existed for millennia, such as Africa.

(Data from: Cook, Earl. "The Flow of Energy in an Industrial Society," *Scientific American* 224, no. 3 [1971]: 135–144; Cook, Earl. *Man, Energy, Society*. San Francisco: W. H. Freeman and Company, 1976; Energy Information Administration. "International Energy Statistics," U.S. Energy Information Administration. http://tonto.eia.doe.gov/cfapps/ipdbproject/iedindex3.cfm?tid=44&pid=45&aid=2&cid=regions,&syid=1980&eyid=2007&unit=MBTUPP: accessed June 10, 2010.)

on fossil fuel energy, which carries with it significant implications for natural systems that are coupled to energy consumption. (For example, the climate system is coupled to the fossil fuel energy system because emission of carbon dioxide from fossil fuel combustion increases the atmosphere's capture of energy radiated from Earth, a process commonly if oversimplistically called **global warming**.)

Even within countries that are clearly on the developed side of the Great Divergence, however, patterns can vary in interesting and significant ways over time. Thus, for example, one finds interesting differences in U.S. energy consumption patterns between 1960 and those of only twenty and forty years later (Figures 1.7, 1.8, and 1.9). Perhaps most obviously, in 1960 coal use measured in **quads** (a quad is 10^{15} BTUs; a BTU is a British thermal unit, which is the amount of heat required to raise the temperature of a pound of water 1 degree Fahrenheit) was only 11 as compared to 12.2 quads of natural gas, but by 2000 coal use was 22.7 quads, not only a doubling of coal use over a forty-year period, but also more than the natural gas use that year, which was only 19.7 quads. Similarly, note that in 1960 transportation used more quads than electricity generation, whereas in 1980 and especially in 2000 the dominance of electricity generation over transportation becomes pronounced. This trend, which has occurred in virtually all developed economies, is known as electrification, and

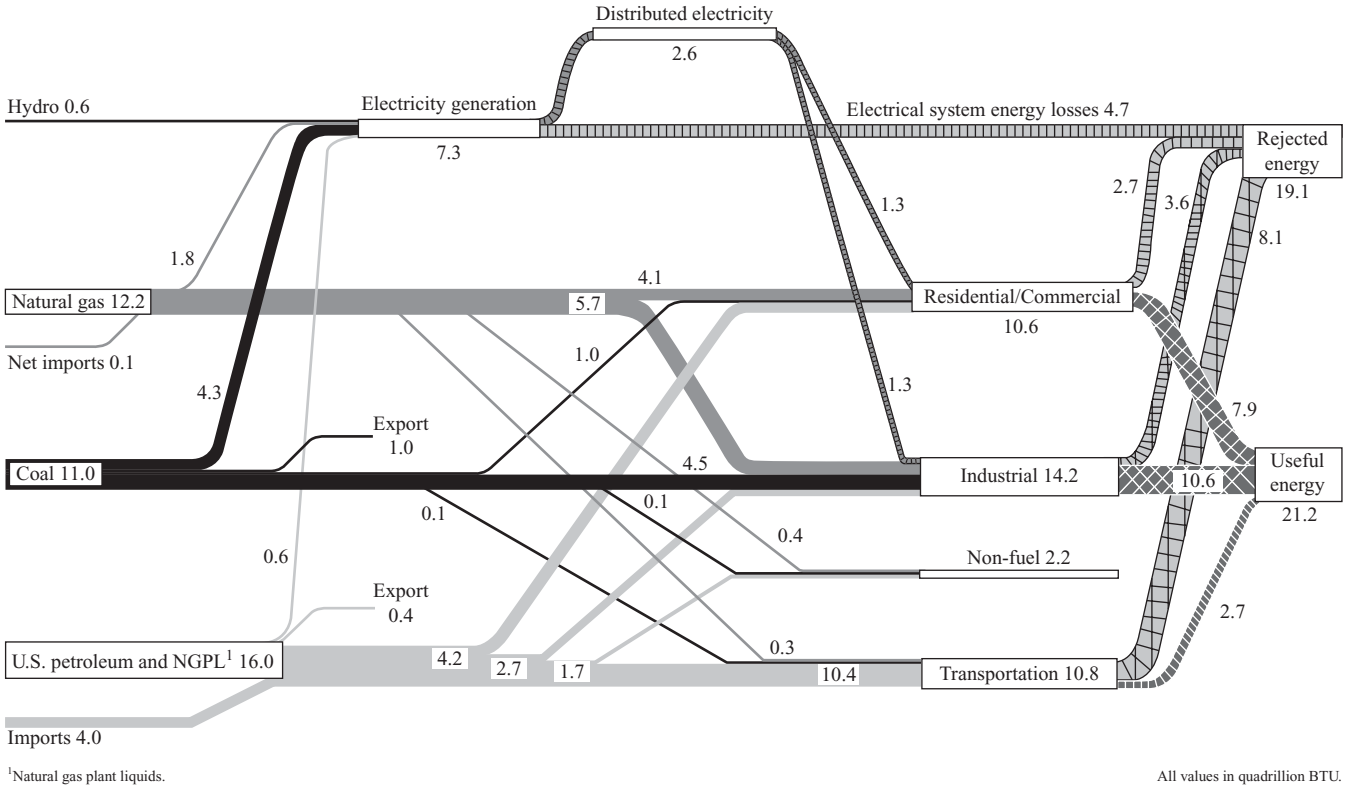


FIGURE 1.7 Sankey Diagram of U.S. Energy Flow Patterns (Circa 1960), Net Primary Resource Consumption 4.39 Quadrillion BTU. Note that less coal than natural gas (in quads) is being consumed, and that only 21.2 quads of useful energy were used in that year.

(Data from: Austin, A. L. and Winter, S. D. "U.S. Energy Flowcharts for 1950, 1960, 1970, 1985, and 1990." https://flowcharts.llnl.gov/content/energy/energy_archive/US_energy_flow_archive/UCRL51487.pdf: accessed January 15, 2011.)

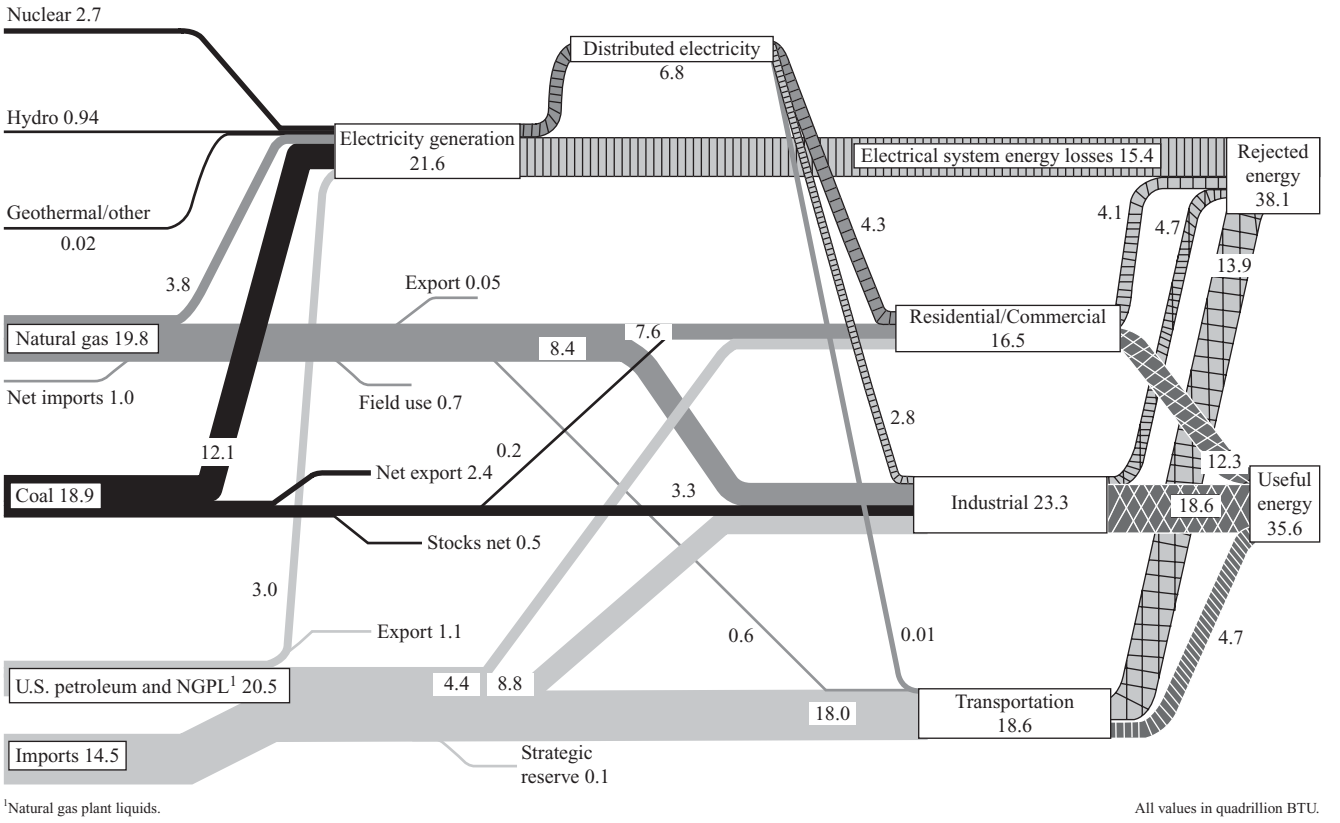


FIGURE 1.8 Sankey Diagram of U.S. Energy Flow Patterns (Circa 1980), Net Primary Resource Consumption 75 Quadrillion BTU. Note the small contribution of "green" energy sources, such as hydro and geothermal, to the mix. The overwhelming sources of energy in 1980 remain natural gas, coal, and petroleum.

(Data from: Briggs, C. K. and I. Y. Borg. *U.S. Energy Flow—1980*. Oak Ridge, TN: Lawrence Livermore National Laboratory, 1981, October 21.)

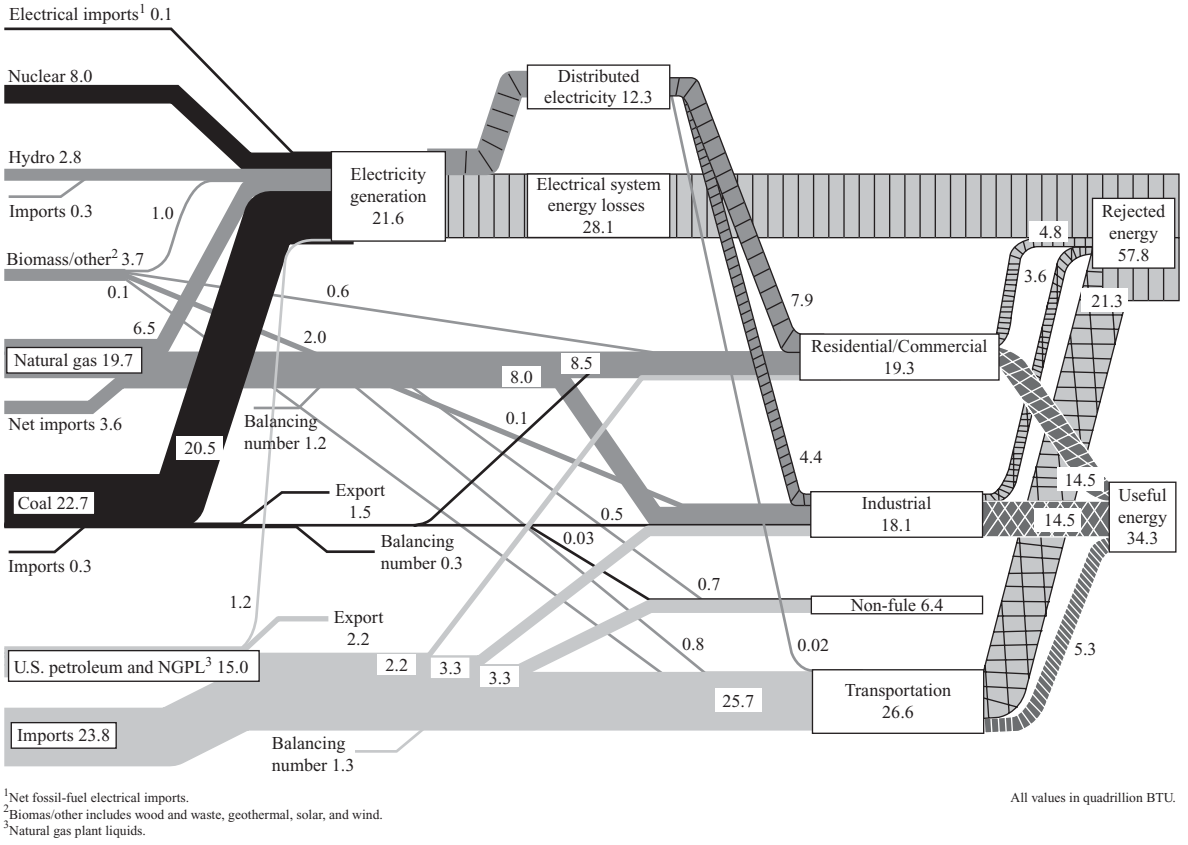


FIGURE 1.9 Sankey Diagram of U.S. Energy Flow Patterns (Circa 2000), Net Primary Resource Consumption 98.5 Quadrillion BTU. The contribution of nuclear energy has grown from earlier years, but remains small; the contribution of biomass and other forms of "green" energy remain insignificant from a systems perspective. "Balancing numbers" represent aggregated figures for minor energy flows involving, e.g., strategic reserves, and are used to reduce the visual complexity of the diagram.

(Data from: Lawrence Livermore National Laboratory and Energy Information Administration. "U.S. Energy Flow Trends—2000." https://flowcharts.llnl.gov/content/energy/energy_archive/energy_flow_2000/USEnFlow00-quads.pdf: accessed January 15, 2011.)

represents a shift away from reliance on primary energy sources, such as coal and petroleum, to secondary sources, especially electricity.

The diagrams in Figures 1.7 through 1.9 illustrate several important substantive points regarding energy shifts over the past half century. Note, for example, that in 1960 the most significant use of energy is in the industrial sectors, followed by residential/commercial and transportation. This was still true in 1980, but by 2000 industrial and residential/commercial had tied for useful energy consumption, and compared to 1980 industrial use of energy, whether measured by input or by useful energy consumed, had dropped by 2000. Thus, the energy mapping reflects the fundamental shift in the U.S. economy towards services and away from manufacturing.

From the perspective of sustainable engineering, there is another, less obvious lesson to take away from these diagrams: the critical requirement that the engineer be able to communicate complex data sets to less technically trained audiences. The value of these diagrams, then, is not just that they capture information about energy systems, but that they present the data even to a casual reader in a way that makes that information intuitively available and intelligible. Because it is ever more important for engineers to be able to communicate complex information to the public and policy makers, this is an important initial lesson to keep in mind.

Patterns similar to those evident at a high level in the energy sectors can be found in water consumption, as Figure 1.10 illustrates. What is important is not

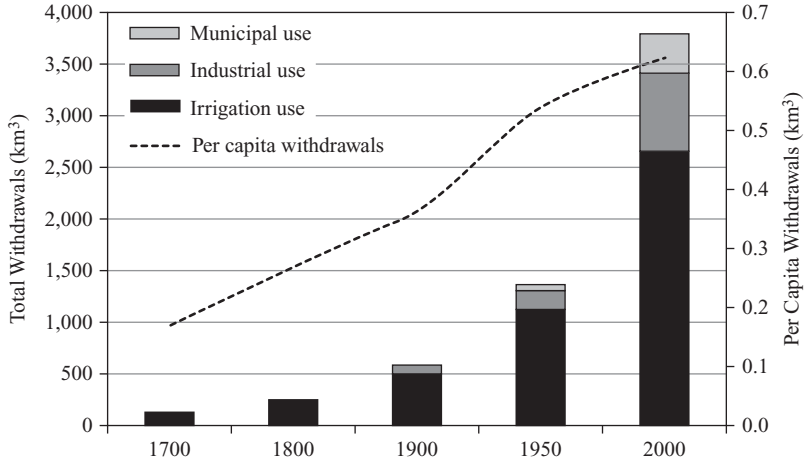


FIGURE 1.10 Global Freshwater Use Over the Past 300 Years. This chart shows a dramatic increase in global freshwater use over the past 300 years, although the general rule that most water is used in agriculture continues to hold over that entire period.

(Data from: McNeill, J. R. *Something New Under the Sun: An Environmental History of the Twentieth-Century World*. New York: W. W. Norton & Company, 2000; World Resources Institute. "Freshwater Resources 2005." http://earthtrends.wri.org/pdf_library/data_tables/wat2_2005.pdf: accessed May 24, 2010.)

just the significant growth in water consumption, which overall has increased by a factor of almost 50 over that period, as well as increasing by a factor of over 4 in per capita terms. It is also important to realize that both industrial and personal use of water are dwarfed by its use in agriculture, especially as the number given counts only the water directly used in irrigation and not indirectly used in agricultural production (for example, the water used to produce fertilizers and pesticides).

While we're on the subject of water, however, let's explore a few important principles that apply to many resources generally. First, looking at data at a global scale can hide trends and developments that have important implications for efficient resource use. Thus, consider Figure 1.11, which shows water use patterns in the United States over the past 100 years. It is apparent from this figure that, despite the consistent per capita increase in water consumption globally that Figure 1.10 documents, water consumption in the United States no longer rises in lockstep with either population growth or economic growth; in technical terms, water consumption has been "decoupled" from population and economic growth, at least for the United States. Knowing why this has happened may enable engineers and policy makers to replicate this trend in other countries, reducing global demand for an increasingly scarce resource, clean water.

But water, like most resources, isn't used just directly by agriculture, industry, or people. Every product we use, or food we eat, has been made using water. This type of trade in water is sometimes called **embedded water**, and it can be very significant, as Figure 1.12 shows. It is noteworthy that two items in

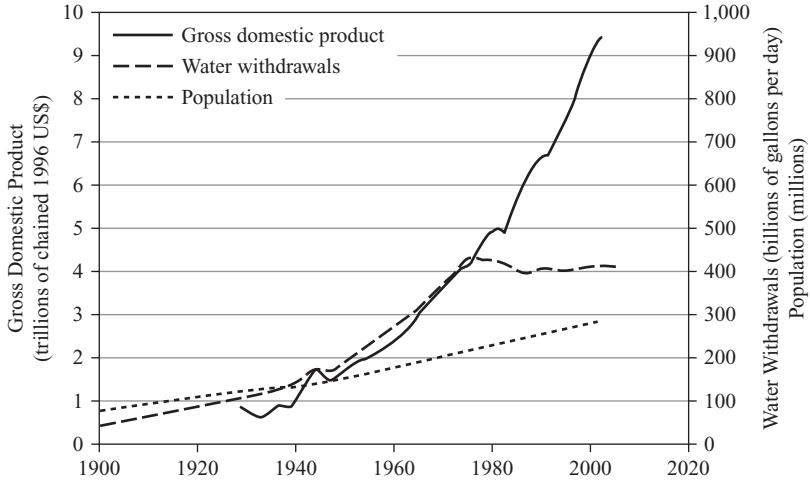


FIGURE 1.11 U.S. GDP, Population, and Water Withdrawals. This graph depicts water consumption in the United States decoupling from both population growth and economic growth.

(Data from: United States Census Bureau [USCB]. "The 2010 Statistical Abstract: Historical Statistics." http://www.census.gov/compendia/statab/hist_stats.html: accessed August 6, 2010; USCB, S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright, eds. *Historical Statistics of the United States on CD-ROM: Colonial Times to 1970*. New York: Cambridge University Press, 1997.)

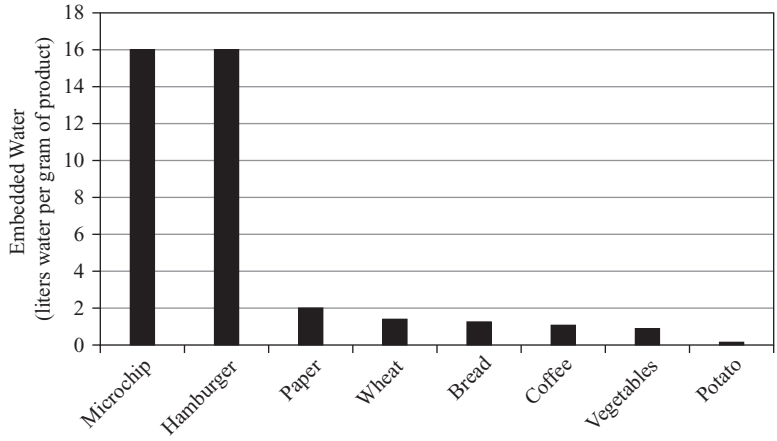


FIGURE 1.12 Embedded Water Content of Selected Items. Direct water use and indirect water use—sometimes called "embedded water"—is important in understanding the flows of materials and resources attributable to products. Such embedded resources can be reduced in many cases by good engineering, but the tradeoffs and design constraints need to be well understood.

(Data from: Hoekstra, A. Y. and A. K. Chapagain. "Water Footprints of Nations: Water Use by People as a Function of their Consumption Pattern," *Water Resources Management* 21 [2007]: 35–48.)

particular contain substantial embedded water: computer chips and hamburgers. In the first case, computer chips require substantial energy to produce, because so much manipulation of matter at very small scales is required to manufacture such items, and energy requires water. An obvious implication for sustainable engineering is that if nanotechnology manufacturing is to become widespread, it will require serious attention to energy and water availability.

With beef, the high embedded-water content is because cows are raised using a lot of fodder, and fodder requires a lot of water to grow. In fact, some arid countries such as Saudi Arabia, Kuwait, Egypt, and Qatar are buying huge pieces of land in poor countries such as Sudan, Cambodia, Ethiopia, and Indonesia so that they can grow food there, enabling them to stretch their scare domestic water supplies much further by "importing water" in the form of food. This "outsourcing of farm production" (sometimes seen as the third wave of outsourcing, after outsourced manufacturing and services) is seen by some as a land grab. Others, however, see it as an efficient way to structure global agriculture, and water demand, and note that such an outsourcing pattern can also provide intellectual and financial capital to poorer countries. Figure 1.13 dramatically illustrates the importance of managing trade to consider the water embedded in food products to an overall global water management regime.

This perspective leads to an "earth systems engineering and management" view of water, in which one begins to map the relationships among the human, natural, and built systems that are engaged by questions of water engineering and policy. Figure 1.14 shows one such mapping for water. The linkages identified in

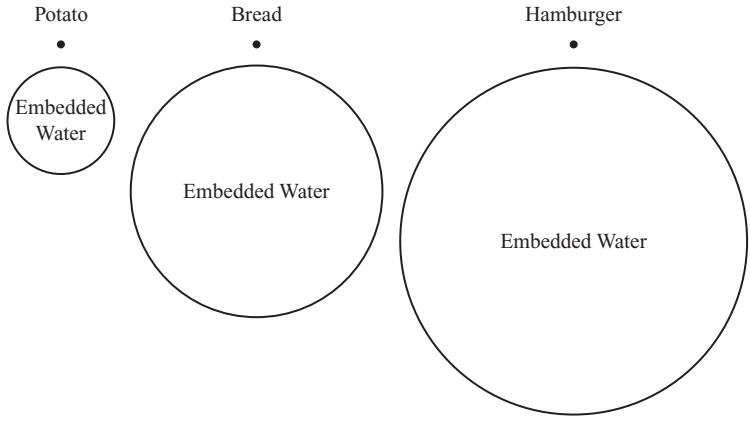


FIGURE 1.13 Embedded Water Content Depicted in a Different Way. Here, the area of the small circles represents the mass of a final product (potato, bread, and hamburger) and the area of the large circles corresponds to the mass of water embedded in the final product.

(Data from: Hoekstra, A. Y. and A. K. Chapagain. "Water Footprints of Nations: Water use by People as a Function of their Consumption Pattern," *Water Resources Management* 21 [2007]: 35–48.)

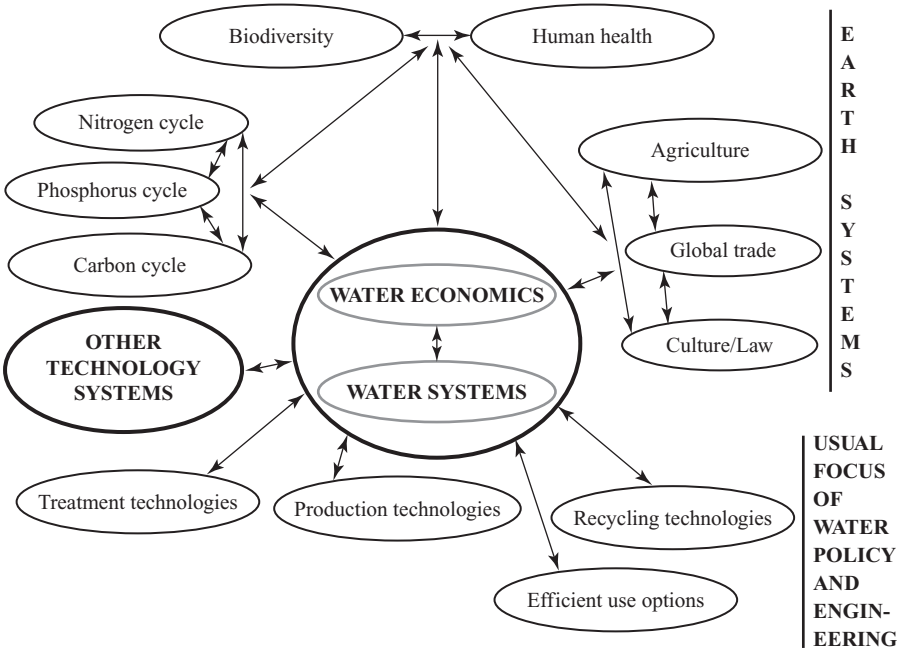


FIGURE 1.14 Water as Earth System. This figure provides a mapping of water as an earth system and differentiates between those aspects of water systems that are usually considered by engineers, such as recycling or production technologies, and those aspects that increasingly must be considered as part of sustainable water engineering, such as applicable laws and the stability of global trade patterns (which can facilitate increased reliance on imported food, and thus imported embedded water, by countries that may lack adequate water resources themselves).

this figure that are important to the sustainable engineer depend on a number of factors, including most importantly the problem presented for solution. If a sustainable engineer, for example, is asked to improve water treatment, it may well be that most of the system map is not highly relevant, because it covers far more than the issues that the problem as stated defines as important (e.g., proper sanitary sewage management, good water treatment plant design). If, however, a sustainable engineer is asked to review the ten- to twenty-year supply of water for a major urban center, a number of issues suggested by the map may be important: Twenty years is plenty of time to try to change outmoded or dysfunctional water laws, for example, or to negotiate new water-leasing arrangements with individual holding water rights, or to change other technology systems, such as manufacturing, to use less water. From a policy perspective, such a mapping can identify opportunities that a traditional water engineering approach might not perceive, such as using new agricultural technologies and patterns of global trade, to increase the efficiency of transfers of embedded water from water-rich regions to countries and areas where explicit water may be scarce.

1.5 THE FINGERPRINT OF THE ANTHROPOCENE: HUMAN IMPRINTS ON SINKS AND MATERIAL FLOWS

These patterns of technological, economic, and population growth, and resulting substantial increases in resource consumption, necessarily increase human impacts on natural systems. The most familiar concern arising from this integration of the human (social, institutional, economic, and cultural systems), the built, and the natural environments is probably "global warming," sometimes called **climate change**. The physics underlying concerns regarding climate change involve a phenomenon called the **greenhouse effect**, which occurs because, just like a greenhouse, the atmosphere lets in light and traps heat. Specifically, high-energy ultraviolet (UV) radiation from the sun travels through the atmosphere, is absorbed by Earth, and is re-emitted as lower frequency infrared radiation (IR). (Figure 1.15 provides a schematic of the physics of climate change.)

The atmosphere neither reflects nor absorbs energy significantly at the UV frequencies (although ozone in the stratosphere does absorb some harmful high-energy UV before it affects life on the surface), but it does absorb the IR fairly efficiently. In particular, water vapor, carbon dioxide, methane, and a few other trace materials in the atmosphere are efficient in absorbing IR. Thus, if the amount of these materials is increased in the atmosphere, the amount of energy absorbed will also increase, with the result that the energy content of the atmosphere is also therefore increased. There is virtually no argument about the basic physics of this process; Venus is a hot planet precisely because of the greenhouse effect (it has an atmosphere rich in compounds that absorb well at the wavelengths radiated by the planet). The debates about climate change—as opposed to the well-demonstrated greenhouse effect—have instead to do with

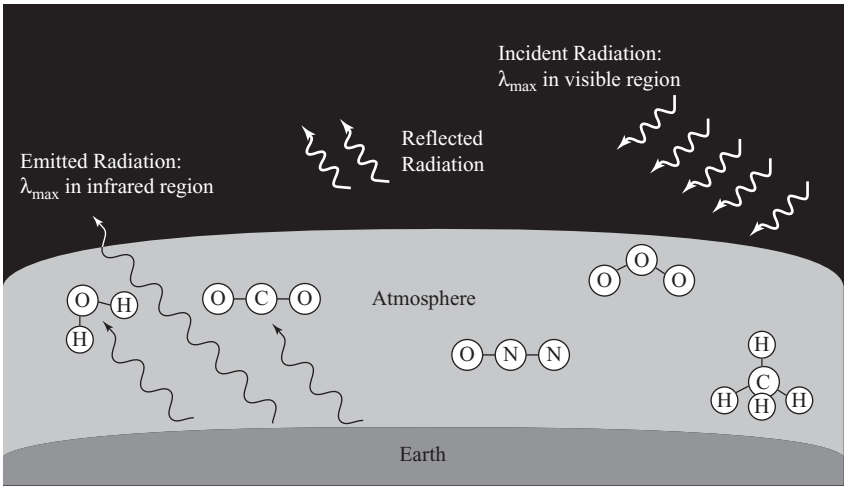


FIGURE 1.15 Earth Energy Model. This figure illustrates the basic physics of the greenhouse effect. Human activity, primarily associated with energy production and agriculture, has over time increased the concentration of chemical species that contribute to climate change forcing in the atmosphere, especially CO_2, CH_4, and N_2O (nitrous oxide).

the magnitude of the effect on global climate from human changes to atmospheric chemistry compared to everything else going on in that complex system, and what should be done on a policy and engineering level in response.

We will discuss global climate change in many places throughout this book. For now, a few basic points should be made. Most important, although global warming has become for most people virtually the only environmental consideration they pay attention to, this is a gross oversimplification of reality and should not be the way sustainable engineers—or any professionals—think about their world. To begin with, climate change reflects only one set of changes that humans are making to the carbon cycle, so it is perhaps more appropriately understood as an aspect of a larger challenge: appropriate management of the carbon cycle. Figure 1.16 provides an overview of the carbon cycle. It is important to emphasize that it is not the absolute magnitude of changes that humans introduce into natural systems, but the effect on the dynamics of the affected system that is important. In this case, the additional volume of carbon that humans are contributing to the atmosphere is small compared with other carbon fluxes, but because it affects atmospheric absorption of energy radiated from Earth, it dramatically affects the dynamics of climate (how dramatically is part of the debate). Moreover, all of these systems are complicated by the fact that they integrate across large physical and temporal scales. Thus, the carbon cycle so simply illustrated in Figure 1.16 goes all the way from annual fluxes to geological processes measured in many tens of thousands, even hundreds of thousands, of years. At the least, this complicates measurements of phenomena;

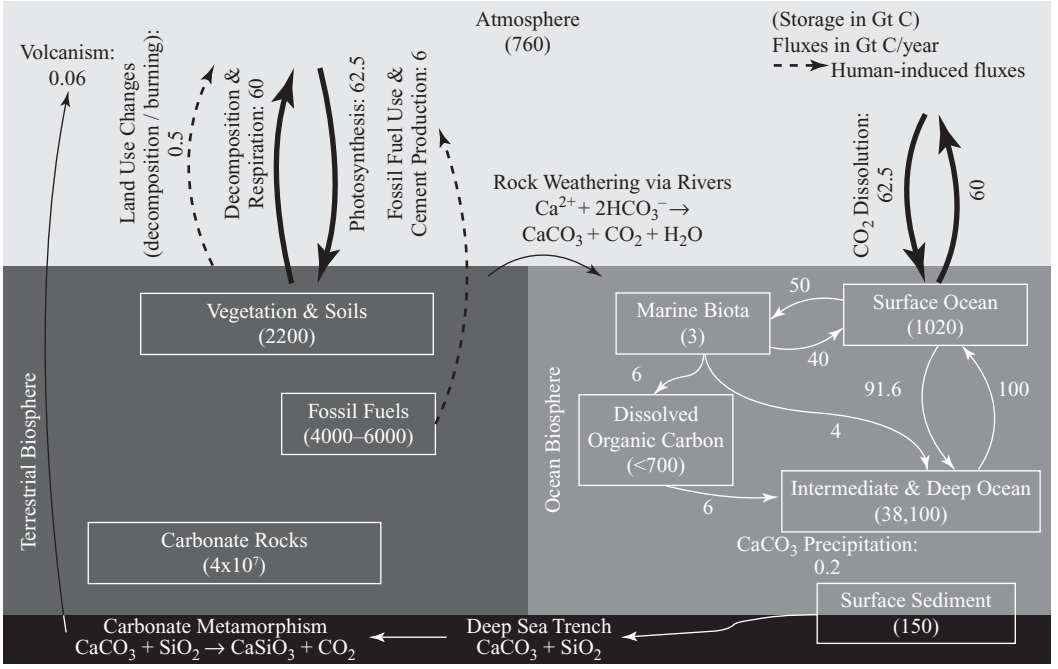


FIGURE 1.16 Carbon Cycle. This figure illustrates some major aspects of the carbon cycle, which is actually an extremely complex combination of networks of effects operating at many different physical and temporal scales.

it also indicates that what we call a "cycle" is actually a very complex set of networks at many various scales. Which of these networks is relevant for a particular analysis is an important, and often difficult, question.

In addition, it is important to remember that the carbon cycle itself is coupled to many other cycles and systems, human, built, and natural. As in Figure 1.14 for water, Figure 1.17 demonstrates that to manage the carbon cycle is to implicitly change many other cycles, from human economic development and technology choices, to other natural cycles such as nitrogen and the hydrologic (water) cycle. It also implies a set of policy and technology choices, which themselves couple to many other systems and choices. In short, what is usually presented by activists as a simple question of **carbon footprint** must be understood by the sustainable engineer in its reality: an extremely complex system that is constantly and unpredictably changing. Among other things, this suggests from the beginning that an important part of the practice of sustainable engineering is being able to simplify and thus understand the system with which the engineer must work, without oversimplifying it. As Figure 1.17 suggests, this is part of the emerging field of earth systems engineering and management, about which we say more in Chapter 14.

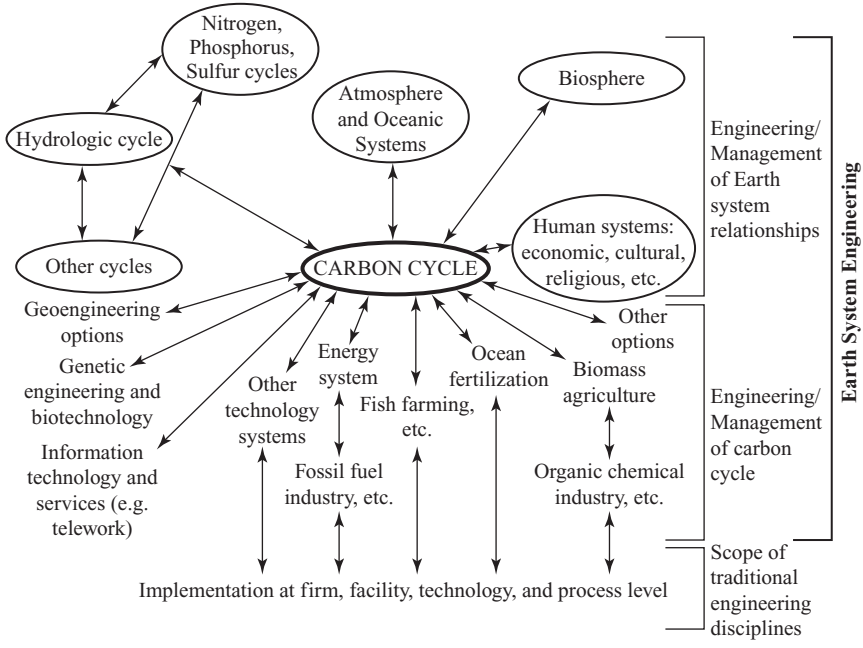


FIGURE 1.17 Earth Systems Engineering and Management: Climate Change–Carbon Cycle Schematic. The carbon cycle is embedded in a framework that captures some of its connections to other systems and issues. This interconnection of natural systems, built systems, and human systems is characteristic of an earth systems engineering and management framework.

Nitrogen offers another example of a "natural" cycle that is heavily influenced by human activities. An element identified only about 200 years ago, nitrogen is a critical component of nucleic acids and proteins, and thus necessary for all life. Although there are substantial amounts of nitrogen in Earth's crust and atmosphere, it cannot be used by biological systems until it is "fixed," or incorporated into reactive compounds. For example, most atmospheric nitrogen is in the form of N_2, and because of the strong bonds in that molecule, it is unavailable for use until split into atoms that then recombine into ammonia or nitrogen oxides of various kinds. Before substantial human involvement with the system, major mechanisms by which nitrogen was fixed included atmospheric phenomena, especially lightning, and biofixation by, for example, legume and soybean systems. But as Figure 1.18 shows, this situation changed fundamentally after the middle of the twentieth century. Although the direct cause was the dramatic increase in fertilizer production and use over that period, contributions from fossil fuel combustion in automobiles and the like were also significant.

The impacts of human changes to natural nitrogen patterns are varied and potentially significant. Some nitrogen compounds have substantial impacts even at trace levels on atmospheric chemistry and physics, in particular by contributing to the depletion of the stratospheric ozone layer or acting as greenhouse gases. Nitrous oxide (N_2O), for example, is about 200 times as potent on a per molecule basis as carbon dioxide in contributing to global climate change. On a more regional basis, anthropogenic nitrogen compounds contribute to photochemical smog and acid rain. Nitrate ions in drinking water can bind to and inactivate hemoglobin in human blood, causing "blue baby"

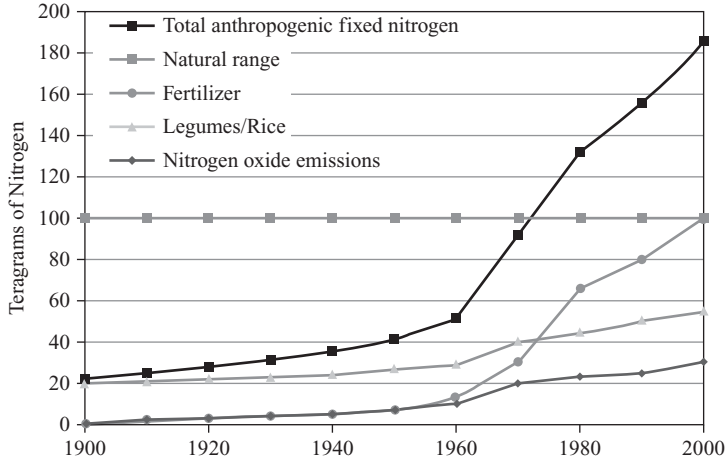


FIGURE 1.18 Terrestrial Fixed Nitrogen. This graph shows human activities affected the nitrogen cycle, especially with the growth in agriculture after World War II: nitrogen fertilizer produced by the Haber-Bosch process was a prerequisite for, and in turn became demanded by, a rapidly growing human population.

syndrome and, at the extreme, infant death. Fertilizers almost never stay just where they are applied, and as fixed nitrogen is frequently a limiting factor in nonagricultural ecosystems, the result of agricultural runoff is the unintentional fertilization of many ecosystems and thus loss of species and shifts in community structure. On a regional basis, the flow of nitrogen from agricultural operations can have significant impacts, especially in estuarial systems. Perhaps the best known example is the **dead zone** that appears in the Gulf of Mexico every year as a result of oxygen depletion on the Gulf floor. This depletion results when the nitrogen from agricultural operations in the U.S. Midwest flows down the Mississippi River to the Gulf of Mexico and causes an algal bloom; the algae and small organisms that feed on the algal bloom then die and sink to the bottom, where their decomposition uses all available oxygen, creating a dead zone. This is not a trivial phenomenon, because the dead zone, and accompanying kill of bottom-dwelling species such as crab and snails, in many years is as large as the state of New Jersey (that is, about 20,000 square kilometers).

The nitrogen cycle, like the carbon cycle, illustrates another important principle. It is a classic case where anthropogenic effects are extremely small in terms of stock, but have significant impact in terms of flows. Eighty percent of the atmosphere is relatively inert nitrogen gas—N_2—and the amount of nitrogen that humans mobilize pales in comparison to the amount in the atmosphere. However, human influence on that small part of the nitrogen cycle involving fixed—biologically active—nitrogen is far more dramatic: in the late eighteenth century, nitrogen was identified as an element; by the middle nineteenth century, nitrogen was recognized as critical for crop production; and by the late twentieth century, humans had grown to dominate the rate of nitrogen fixation. As in many cases, the important aspect of a material flow is not necessarily its absolute magnitude, but its effects even at small concentrations.

The extent to which humans now dominate a number of "natural" material flows is remarkable. Table 1.2 illustrates this phenomenon for a number of elements; for some, such as calcium, the human domain is relatively small, but for many others—nickel, lead, iron, copper, and zinc, for example—the amount of the material mobilized naturally is less than the amount that humans mobilize. The effects of the activity required to produce these materials are, of course, not trivial (e.g., mining, shipment of ores). But an engineer must also consider stability of supply, and availability, especially when dealing with designs that may be implemented at large scales (e.g., an important part for automobiles may not require much material individually, but because automobiles are produced by the millions, total demand may be significant). Particularly with rarer materials, such as bismuth and indium, for example, resource availability can be an issue.

Figure 1.19 shows the environmental side of the economic and technological story told earlier. Systems that had been stable for thousands of years,

TABLE 1.2 Human Mobilization of Selected Elements[1]

Element	Total Mobilized (Tg/year)	Anthropogenic vs. Natural Ratio	Anthropogenic as Percent of Total
Arsenic	0.1	1.7	63
Carbon	118,450	0.1	9
Calcium	7,249	0.05	4
Chlorine	6,264	0.02	2
Copper	16	5.9	85
Iron	848	8.9	90
Nitrogen	6,067	0.1	7
Nickel	2.3	2.3	69
Lead	3.9	5.2	84
Sulfur	766	0.3	22
Zinc	20	0.9	47

[1]This table indicates that the cycles and flows of many elements, metals and nonmetals, are increasingly a product of human decisions, and economic and technological imperatives, rather than natural forces. (Data from: Klee, R. J. and T. E. Graedel. "Elemental Cycles: A Status Report on Human Or Natural Dominance," *Annual Review of Environment and Resources* 29 [2004]: 69–108. Reprinted with permission.)

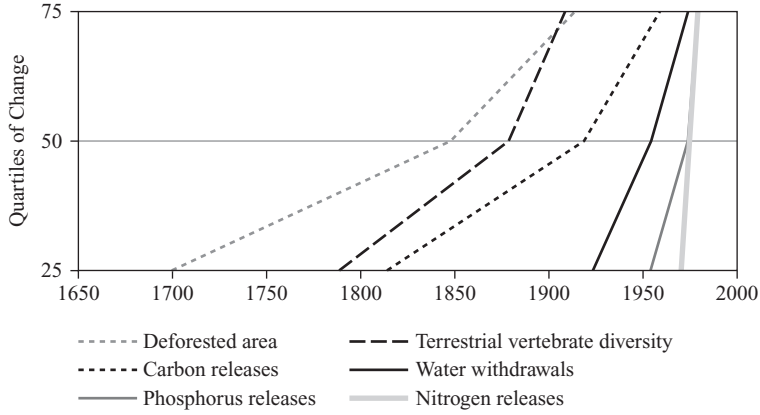


FIGURE 1.19 Human-Induced Environmental Processes. This graph is composed by setting the far past (circa 10,000 years ago) as 0% change, and 1985 to be 100% of the indicated change. It then illustrates when each phenomenon reached 25% of its current value, and when it reached 75%. Each process, from deforestation to carbon and nitrogen releases, was relatively stable until the Industrial Revolution, then accelerated dramatically. This acceleration is the environmental side of the population, economic, technological, and material growth illustrated previously.

(Adapted from: Turner, B. L., W. C. Clark, R. W. Kates, J. F. Richards, J. T. Mathews, and W. B. Meyer, eds. *The Earth as Transformed by Human Action: Global and Regional Changes in the Biosphere Over the Past 300 Years*. Cambridge: © Cambridge University Press, 1990. Reprinted with permission.)

from biological community diversity to deforestation, underwent dramatic change. This should not, however, be taken to imply a Golden Age when humans were "in harmony with their planet"; there is no indication that any particular culture was better in that sense than any other (in spite of our tendency to romanticize other cultures as more "sustainable"). Rather, all indications are that the difference between primitive communities with less impact, and modern society, is simply that the former didn't have the technology to do more damage than they did. What Figure 1.19 illustrates, then, is not some fundamental change in psychology, but rather the impacts across many systems, natural and human, of the extraordinary economic, demographic, technological, and social changes that characterized the Industrial Revolution. At this stage in our evolution on this planet, we cannot separate the engineering, management, and policy decisions into neat categories, nor can we bound our technology decisions as simply as, for example, the popularity of "carbon footprints" or even **environmental footprints** would suggest. The sustainable engineer must be far more sophisticated than that, because the effects of engineering and technology in today's world require it. An example is provided in Box 1.3, which discusses the case study of corn-based ethanol.

BOX 1.3 Case Study: Corn-Based Ethanol

From the perspective of global climate change, plants have the important property of capturing atmospheric carbon as they grow and thus (potentially) are "carbon neutral" when used to produce products that substitute for fossil fuels (e.g., biodiesel or ethanol/gasoline mixtures). Accordingly, in 2001 the U.S. government explicitly adopted the goal of tripling the use of "biobased products and bioenergy" by 2010, establishing a Biomass Research and Development Board (BRDE). The BRDE issued a report that, among other things, declared that one of the goals of the U.S. biofuels effort was environmental improvement, because it was, according to the BRDE, "common knowledge" that biomass is a renewable source of carbonaceous material and energy, and is therefore preferable to alternatives. Thus, the Bush administration in 2008 gave preferential treatment to domestic, corn-based ethanol by imposing a 54-cent-per-gallon tax on imported ethanol (aimed explicitly at the more efficient Brazilian producers of ethanol from sugar cane), as well as a 51-cent exemption from the federal excise tax on gasoline for fuel mixed with ethanol.

Unfortunately, it takes a lot of energy to raise corn, and even more to turn it into ethanol at the scale required by the U.S. initiative—and most of that energy comes from fossil fuels. Thus, depending on how the analysis is structured, it is likely that little if any carbon emissions are saved by switching to corn-based ethanol. Moreover, as the figures in the text demonstrate, once a technology such as this is implemented at any significant scale, a lot of other systems will be perturbed. In this

case, an easily predictable perturbation was that food prices rose significantly as corn was shifted from food use (either directly as a grain, or indirectly to raise livestock) to corn ethanol production. As prices rose, there were riots in a number of countries, including Mexico and the Philippines, by poor people who were forced to go hungry by rises in the price of basic food commodities. Moreover, farmers are adept at detecting fundamental market shifts, and thus began shifting their land from other crops into corn, which meant that food prices across the board began to increase. Longer term, if demand for agricultural products permanently increases because of biofuel production, it implies use of land not now in agricultural production, with ripple effects on biodiversity and other systems affected by land use. Increasing demand for agricultural products and their embedded water, leads to countries purchasing large pieces of land in other, usually poorer countries, to maintain production. Therefore, more ripple effects are created.

But it is even more complicated than that. After all, biomass is not just carbon, as the focus on climate change mitigation tends to imply; it includes nitrogen, phosphorus, and a host of other elements as well. Thus, expansion of agriculture for purposes of producing biomass may be carbon neutral under ideal conditions, but it is not nitrogen neutral, as Figure B1.3.1 illustrates. And, as discussed in the text, the nitrogen cycle is already deeply perturbed. The last thing it needs is a large spike in use.

From the perspective of sustainable engineering, this cautionary tale should not be taken to mean that biomass energy and material production is not potentially a worthwhile technology. Rather, it suggests that, before introducing a large-scale new technology into a globalized and interconnected world, at least some preliminary analysis of potential benefits and costs is useful—indeed, it is a professional imperative. More specifically, this case also illustrates that it is critical in evaluating these sorts of decisions to understand when important variables are liable to go seriously nonlinear, thus producing unanticipated and potentially significant unintended consequences.

This example also illustrates another point. Many of the systems that sustainable engineering deals with, such as the corn-based ethanol biofuel system, are known as **complex adaptive systems**; they are very complicated, they shift over time unpredictably, and they are usually tightly coupled to not just other technological systems, but also to social and environmental systems. In these cases, one needs to be very careful to include important and relevant results of taking a particular action or implementing a particular technology. It is doubtful, for example, that the United States intended to destabilize close allies by causing food riots. On the other hand, it is also relevant to this technology that a major driver for its rapid adoption at large scale was political: Farm state senators in the United States are politically very powerful, and this policy was a good way of getting their support for other political initiatives having little to do with biofuel. These kinds of complications put a substantial burden on the sustainable engineer, but increasingly they cannot be avoided.

In addition, note that the problems with the corn-based ethanol arose not from any particular locus—the inherent dynamics of the carbon or nitrogen cycles,

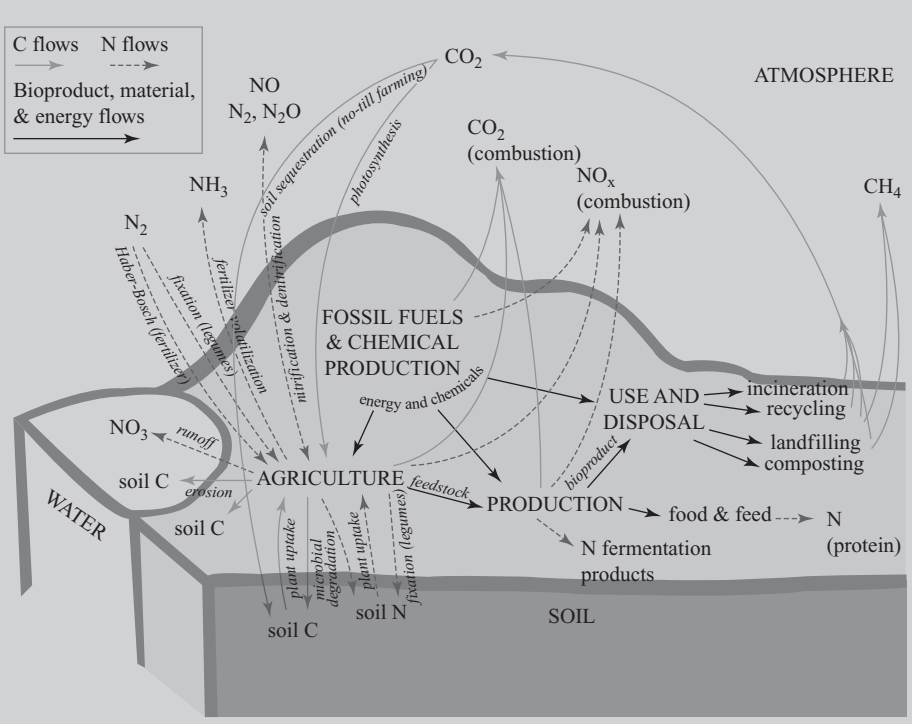


FIGURE B1.3.1 Lifecycle Carbon and Nitrogen Fluxes for Bio-Based Products. This image illustrates some of the interconnections between the carbon and nitrogen cycles that become important as one begins considering biobased products such as corn ethanol as a biofuel. Studies of such interconnections should precede, rather than follow, implementation of major technology initiatives.

for example, or the politics, or the various technologies at issue—but rather emerged from the complex system as it evolved through time. Although some emergent behaviors should be relatively easy to predict, such as higher food prices, others (e.g., riots, and outsourcing of agriculture to poorer countries) are less obvious. These so-called emergent characteristics or **emergent behaviors** emerge from the system as a whole in real time, and thus rational design and management of even subsystem elements (e.g., biomass as a partial solution for global climate change forcing) require some knowledge and sensitivity to the system as a whole. The existence of often-acknowledged "unintended consequences" is, in fact, an important indication of the gap between efforts to rationally design and manage subsystems, and the consequent responses of the overall system.

Further, the engineering professional needs to be aware that even policies claiming a basis in "science," especially in the environmental and sustainability areas, are usually highly subjective. The policies will frequently combine ideology, myth, and lack of data into solutions that fail to respect, or even consider, the scale and complexity of the relevant systems. Corn-based ethanol appealed to many people because everyone "knew" that "biomass" is "green" and "industrial energy

infrastructure" is not. This knowledge, of course, was not based on science and analysis of the system in question, but rather on environmental ideology. Few real-world systems with which engineers must grapple are quite that simple, and the more "everyone knows" such things, the more the engineer should be on her or his guard.

Finally, of course, it is noteworthy that human technological and economic systems, in this case involving the energy and agricultural sectors, are often the mechanisms by which system coupling occurs, not just between relevant human systems, but also between "natural" systems—in this case, the carbon and the nitrogen cycles. To understand the Anthropocene without understanding technology is, in short, impossible.

1.6 CONCLUSION

So in what essential ways does our present situation differ from the past? Simply put, our technologies and economies are much more powerful. They enable us to live in numbers, and to capture and control energy, materials, and biological services, in far greater amounts, than we ever could in any previous period of history. Moreover, as subsequent chapters will detail more fully, the context such a world creates for engineering is far more challenging and complicated than anything that has gone before, requiring a profoundly different approach to engineering. Engineers are increasingly called on not just to solve technological problems, but to help society understand the technologies that increasingly define our lives, and to maintain social, economic, and environmental stability in the midst of constant change. The challenge of sustainable engineering is thus easily framed: It is engineering rationally, ethically, and responsibly in an increasingly complex and unpredictable world characterized by accelerating technological change, where myriad human, economic, technological, built, and natural systems are coupled together in complicated ways at regional and global scales.

Engineering students are already taught many of the requisite skills to respond to these challenges. However, surveys of employers of engineers from many universities yield strikingly similar results: most engineers are technically competent, but they lack communications ability and they do not understand the context within which they are expected to perform professionally. This book is an attempt to remedy these deficiencies; if you read it, and by and large understand it, it should give you a significant edge in understanding the manifold challenges of engineering in an increasingly complex world. Moreover, the homework exercises, unlike almost all the other homework you will find in your engineering courses, are not quantitative. Rather, they are intended to address communications skills, and thus are in the form of short written essays

(Box 1.4 provides some general guidelines for effective communication and writing). Do not fear: You will get plenty of practice with mathematical problems in other engineering courses. But you may not get as much writing experience as you need and as your future employers would like to see so hopefully you will benefit from, and even enjoy, these exercises.

BOX 1.4 Writing for Your Audience

Informal surveys of potential employers consistently shows that college graduates, especially engineering students, tend to lack good writing skills. These skills are important to obtain good jobs and for advancement once you are in an organization. They are also critical because they help you get your points across, whether in a proposal to a client or customer, or in a response to your boss. Although good writing is learned over many years, and is the subject of entire courses, a few simple hints can help you improve your writing (for your homework in this course, for example).

1. Be aware of what you are writing and for whom you are writing it. Some writing, such as an informal email to your boss, may not require much editing or formal language. Some writing, such as the homework essays in this book, will require good writing skills and attention to spelling and grammar, but will generally not require citations and footnotes. Other writing, such as a research or term paper, will generally require specific formats, citations, and writing styles; check with your professor to see which one will be required. If you are writing a formal document, such as a contract proposal, see if you can obtain a model to follow. If there is no specific model, choose a style yourself and follow it carefully. The Internet can provide many resources for research-specific styles, such as the APA (American Psychological Association) style often used for social science writing.
2. Be careful how you present quantitative data, especially when you are writing for a general audience. Do not use numbers where they may mislead readers into a false sense of certainty, remembering that many nonengineers and policy audiences tend to presume numbers are concrete and precise just by virtue of being numbers. Where you do use numbers, only provide the precision that your equipment, calculation methods, and data support: Each digit, including the most precise one to the furthest right, should reflect something about the real world. Technically, this means that you should present only "significant digits," especially when rounding numbers. For example, if you are giving the population of Arizona in 2008 in a book, it is better to use 6,500,000, with two significant digits, rather than 6,509,923, with seven significant digits, because the latter implies that you know that in 2008, there were exactly that number of people in Arizona (you never know exactly how many people are in a large region, and even if you did, the obvious problem is

that there would be significant variations over the period of a year). Where you do present numbers that may necessarily be somewhat uncertain—for example, estimates derived from a complex model such as the global circulation models (GCMs) used in climate change research—make sure you convey the uncertainty in the accompanying text, and where possible be clear about the simplifying assumptions you have made in building your model.

3. Think of any writing assignment as requiring three pieces. First, there is an introduction, which should give the reader a good idea of what you're going to talk about, without too much detail. For a one-page essay, this might only be two sentences; for a research paper, it might involve an abstract (a special section summarizing the entire paper); for a contract or memorandum, it might involve a brief summary of the situation as well. Second, there is the body of the essay or paper; this should be well organized so that the reader can follow your points easily, and should not wander or bring in unnecessary points. Depending on the style you are using, footnotes can be used to explain details that would interfere with the smooth flow of your text. Third, there is the summary. Like the introduction, this may only be a sentence or two in an essay; it may be several pages in a long research paper. The summary should reinforce your introduction, without simply repeating it.
4. Even with a short essay, prepare an outline. Even people with good writing skills often produce poor assignments because their work is disorganized, doesn't flow well, and confuses the reader. You don't want your reader to finish your essay wondering what it was supposed to be about. With most writing assignments, a good rule of thumb is to spend half your time doing the outline and half doing the actual writing; if you do the first well, you will find that the second will almost do itself.
5. Pay attention to citations. You won't need them for informal writing, or for most essays, but you will need them for research or term papers. Using the right citation style is of course important, but it's also important to know what you should use as references. Don't use Wikipedia; it is a good place to start trying to understand something, but because entries can be (and sometimes are) modified by people with strong interests in the topic, it is not a good source to cite in a paper. "Peer-reviewed" sources—that is, publications that ask experts to review submitted articles to make sure they're accurate—are preferred as sources, although other sources can be used if appropriate. Whenever you actually quote someone or use text from a particular source, you should always provide the citation: Not to do so can be considered plagiarism, which is unethical and at some schools can lead to disciplinary action.
6. Read your essay aloud to yourself when you are through. In many cases, this will help you catch common mistakes, such as mismatches between verb and subject, and incomplete sentences.
7. If you have time, and the situation allows it, have a friend edit your work. Often, that person will identify problems you weren't aware of and, equally important, he or she may be able to spot sentences that you thought were

clear but that may not be clear to the casual reader. If you are doing this for a homework or class assignment, check with your professor to make sure it's okay to work together on the assignment.

8. Do not plagiarize. If you take text or information directly from another source, make sure you cite that source and, if appropriate, put quotes around the language you are using. In some cultures, it is considered acceptable to copy blocks of text in homework assignments and elsewhere, but in the Western scientific and technological tradition, plagiarism—claiming someone else's work as your own—is considered to be a serious offense in any sort of professional and academic writing, and may be severely punished by your institution. Moreover, be aware that modern grading systems provide the ability to run an essay against virtually the entire content of the Web to make sure it wasn't pulled from an online source; you are unlikely to get away with such a trick for long. The same is true with any publication or article; plagiarism in such documents is considered serious professional malpractice. If you have any questions at all in this area, speak to your professor or instructor, or an institutional advisor, because the results of getting this wrong can be substantial.

Exercises

1.1 Figure 1.9 shows water consumption in the United States going down over the past one hundred years despite population and economic growth. Why might this be the case? Formulate a hypothesis—a testable proposition—based on your speculation.

a. What information and data would you need to determine if your hypothesis is correct?

b. What information and data would you need to determine if you could effectively transfer the technologies and policies that might be the cause of this water consumption reduction to countries other than the United States?

1.2 Human population growth appears to be roughly correlated with the technology systems that are available: Agricultural societies supported more people than hunter–gatherer societies, and industrial societies can support more people than agricultural societies. Why?

1.3 Studies show that the people who first lived in agricultural settlements in the Middle East were generally unhealthier, and apparently died younger, than those who remained in hunter–gatherer societies. Why, then, did anyone move from hunter–gatherer social systems to agricultural societies?

1.4 Urban geographers such as Jane Jacobs and Ed Soja make the claim that hunter-gatherers formed dense settlements in places such as Jericho and Catal Huyuk, before forming agricultural societies, and that the synergy afforded by humans living in proximity was one of the drivers for the very innovation of agriculture.

If you accept their thesis, what are the likely reasons for the subsequent development of larger urban settlements? Where, for instance, did farmers trade their produce and how did people ensure that products other than food were widely available in prehistoric times?

Annotated Bibliography

1. J. R. McNeill, 2000, *Something New Under the Sun* (New York: W. W. Norton & Co.), and B. L. Turner, W. C. Clark, R. W. Kates, J. F. Richards, J. T. Mathews, and W. B. Meyers, eds., 1990, *The Earth as Transformed by Human Action* (Cambridge: Cambridge University Press). Most of us, including engineers and those who work in the sustainability field, frequently lack a sense of history and, accordingly, are not very good at explaining what, exactly, has changed over time to elevate sustainability issues to the level of concern they now occupy. The short answer is "scale," but these books provide a far better background than that. It is often the case, by the way, that edited books are of more uneven quality, and sometimes less coherent, than books that are written by only one author, a general rule that is true in this case.
2. K. S. Robinson, *Red Mars* (1993); *Green Mars* (1994); and *Blue Mars* (1996), (New York: Bantam Books). This series, like many other science fiction books, deals in a fictional way with the ethical implications of terraforming a planet. What is interesting is that there is a large and relatively sophisticated literature on the ethical implications of terraforming other planets, especially Mars, but virtually no discussion of those implications in the case of the only planet that, so far as we know, has been in fact terraformed—Earth. For those who are interested in the subject of terraforming in a nonfictional sense, there is an increasing number of options, as more writers address the possibility. Still, remember that Earth is our first shot, and it's a pretty cautionary one at this point.
3. V. Smil, 1997, *Cycles of Life: Civilization and the Biosphere* (New York: Scientific American Library). Vaclav Smil is a good guide to the anthropogenic Earth, and this book is a good, and very accessible, discussion. Importantly, Smil does not play down the interconnections among the human, built, and natural systems at many different scales that characterize the planet, which makes this a good initial entrée into the integrated systems that are the context for the practice of modern engineering.
4. Good books on terraforming are hard to find. One of the best early efforts, which is hard to find because it's out of print, is M. J. Fogg, 1995 (originally published in 1960), *Terraforming: Engineering Planetary Environments,* which is published by, of all people, the Society of Automotive Engineers. (Do they know something we don't? Is that what automotive technology is about?) A more recent and simpler overview can be found in M. Beech, 2009, *Terraforming: The Creating of Habitable World* (New York: Springer).

CHAPTER

2 Themes of the Anthropocene

So long as we do not, through thinking, experience what is, we can never belong to what will be. . . . The flight into tradition, out of a combination of humility and presumption, can bring about nothing in itself other than self deception and blindness in relation to the historical moment.

—Martin Heidegger, German philosopher[1]

2.1 INTRODUCTION

That we now live in the Anthropocene, on an anthropogenic planet, reflects the reality that our Earth is increasingly shaped by human activities. This concept doesn't mean that the entire Earth is artificial, as it clearly isn't. What it does mean, however, is that the Earth as it now exists increasingly reflects human activities. Thus, for example, Switzerland as a nation appears quite wild in places, especially in the Alps, but it is a profoundly human landscape: Cows graze on the steepest slopes and humans, the ultimate predator on this planet, have long since driven other predators in the Swiss Alps to local extinction (the last

[1]M. Heidegger, *The Question Concerning Technology and Other Essays*, tr. W. Lovitt (New York, Harper Torchbooks, 1977), "The Turning," p. 49; "The Age of the World Picture," p. 136.

bear in the Swiss Alps was killed in 1905). Similarly, we now know that the Amazon was heavily engineered before Europeans even arrived on the South American continent: soils were heavily modified for productivity by indigenous tribes, remains of significant earthworks extend throughout the Amazon basin, and large urban centers existed in the jungle. So is the Amazon "natural"? Of course, and it would be folly to pretend that we understand the complex ecosystems of that region in detail. On the other hand, is the Amazon "anthropogenic," a product of human activity, of human design? That is also true.

As Chapter 1 pointed out, it is not simply by chance that the world appears to be experiencing a dramatic increase in challenges to social and environmental stability. Each individual contributor—more people, more economic activity, more travel to sensitive environments—may be easy to identify and understand, but in the real world, these issues and activities are not so nicely disconnected. In fact, most of them are coupled to each other in very complicated ways, and it is the behavior of this interconnected system that is so challenging—challenging not just to understand, but to the engineer, as it forms the context within which intentional changes directed towards specified ends must be designed and must work.

To discuss all aspects of the anthropogenic Earth would be not just confusing, but impossible. Moreover, some, such as the dramatic implications of emerging technologies, will get their own chapter later. But there are four themes that are important to highlight. These themes will be introduced briefly, because they will affect not just the way engineering will be done in future, but also the careers that engineers will face. These four are globalization and multiculturalism; the changing role of information and information systems; technology's importance for issues of environmental quality and sustainability; and demographics, the latter including urbanization, age-related issues, and increasing risk and vulnerability for infrastructure and human systems resulting from geographic population shifts.

2.2 GLOBALIZATION AND MULTICULTURALISM, THEN AND NOW

Most people these days are familiar with globalization, the continuing integration of social, political, economic, industrial, technological, and information systems around the world. But few realize that globalization in some form, especially in terms of formalized trade, has long been the rule, not the exception, although it has not always been the European West that has dominated that behavior. Trade routes knitted together the old empires of the world—the Roman, the Persian. As Figure 2.1 shows, a thousand years ago the world was already linked by established patterns of trade; however, these patterns were

From Region	To Region: Europe	Islamic World	Central Asia	Sub-Saharan Africa	South Asia	Southeast Asia	East Asia
Europe		Slaves, swords, furs, silver	Furs, swords				
Islamic World	Pepper, spices, textiles, silk, silver		Textiles	Salt, textiles, manufactures, swords, horses	War horses	Gold	Spices
Central Asia	Silver	Paper, silver, slaves			Silver, re-exports from China and Muslim world		Horses
Sub-Saharan Africa		Gold, slaves, ivory, rice			Timber, iron		
South Asia		Pepper, spices, silk, teak, textiles	Pepper, textiles	Textiles		Textiles, Pepper	Textiles
Southeast Asia		Spices, perfumes			Silk, spices, teak, rice, rubies		Perfumes, spices, sandalwood
East Asia		Silk, porcelain	Silk, tea		Silk, porcelain	Silk, copper, cash	

FIGURE 2.1 Interregional Trade Flows, 1000 BP. This table shows the products traded across established routes circa 1,000 years ago between the major regions of the world, with the Islamic World dominating global trade at the time. The obvious exception is the Americas, which, separated by the Atlantic and Pacific from Europe and Asia, were not part of this ancient trade system.

(Information from: Findlay, R. and K. H. O'Rourke. *Power and Plenty: Trade, War, and the World Economy in the Second Millennium*. Princeton, NJ: © Princeton University Press, 2007. Reprinted with permission.)

dominated by the Islamic World, which at that point was trading with every other region of the world, from China and Southeast Asia, to Africa, to Europe. In fact, at that time the Islamic World was a sophisticated, intellectually curious, and productive culture, while Europe was a relative backwater that dealt primarily in slaves and relatively unsophisticated products. In addition, this was not a truly global trade network: North and South America were not part of it because of their oceanic isolation from the Eurasian land mass.

In terms of ship technology, it was China that dominated this world; Chinese junks were far more seaworthy than their European counterparts. As late as 1420, the Ming navy consisted of more than 3,800 ships, 1,350 of which were designed for combat, with 400 of those being large "floating fortresses."[2] Admiral Cheng Ho's expeditions around the Indian Ocean of 1405–1433 involved hundreds of ships and tens of thousands of sailors and soldiers, and dwarfed anything Europe could dream of, both in scale and in engineering competence. But in one of the most significant decisions in world history, the Chinese deliberately turned their back on exploration and expedition; in 1433 naval activity in the Indian Ocean was terminated and a decree was issued in 1436 forbidding construction of any seagoing ship. There is much discussion in history books about why such a decision was taken; counterfactual history should always be taken with many grains of salt, but given the dramatic lead in technology and capability that the Chinese had at that time, it is highly likely that, had they not voluntarily destroyed their oceanic fleet, they would have been the world's dominant power for the next millennium.

But China so decided, and by the later 1400s, Portugal had developed the caravel, the first European ship that could mount cannon (and thus both protect and force trade), sail anywhere, and carry enough to make trade profitable. The Chinese junks of that time, had they been fielded, would have given caravel technology a run for its money, but the Chinese at that point were letting their fleet rot; not only had they stopped expanding their oceanic trade routes, but they were pulling back to China proper. Reasons are not clear, but include a Mongol threat to their northern borders, a desire for internal security and stability rather than expansion, domestic political intrigues that those favoring expansionism lost, a reduction in the financial resources available to the Chinese state, and a desire to focus development resources internally rather than in external expansion. However, even then the world was also linked by information systems and technology transfer; it is quite possible that the Portuguese learned from the Chinese, and highly probable that they learned a lot from the Islamic World, especially regarding sail design and navigation techniques. This latter observation is critical, because it makes a point about all technology: It is not enough to

[2]These data are from William McNeill's 1982 book, *The Pursuit of Power* (Chicago: University of Chicago Press).

build the artifact. The "soft" side of the technology—navigation techniques, the know-how to build an oceangoing ship, the way different technologies are integrated to work together, the technology and trained seamen necessary to actually sail the ship—all this must also be learned, or developed.

The Portuguese, however, were a small country, and were rapidly displaced by the Spanish and, surprisingly, the Dutch. Although The Netherlands was also a small country, the people of this country learned to build cheap ships and, even more important, how to finance them: It was "cheap money" and innovation in banking (driven in part by Calvinism), not just or even primarily technological advantage, that fed Dutch trade ascendancy in the late 1500s (a historical example of the principle that finance and technology are important partners). And when the English subsequently grew to dominate world trade during the period of European imperialism, beginning in the 1660s and lasting almost until World War I, it was again not just technology, but a series of technological, institutional, and social innovations that drove their ascendancy. In particular, the English early on developed financial mechanisms that gave the English nation-state a dominance that no other European society (particularly the French) matched, and they were more effective at manning their ships, both military and civilian, because they were more effective at moving agricultural workers off the land, from whence they had to find occupation somewhere else.

English domination of world trade was part and parcel of the Industrial Revolution, which, beginning in the United Kingdom, spread both to English colonies such as those in North America, as well as to other European powers (although it was only at the beginning of the twentieth century that Germany, and on the other side of the Atlantic, the United States, began to catch up with, and in some areas overtake, English manufacturing capability). Nonetheless, despite the English navy and its dominant international role, other European countries, especially the French, continued to compete in global trade. Culture rode the back of trade, colonialism, and financial success, with the result that Enlightenment values and institutions, such as the nation-state, became globally dominant. But even though this European trajectory is perhaps the most familiar to most readers, it should always be remembered that until perhaps 1800 the largest national economy was not European, but Chinese, and, given current events and shifts in economic power, it may soon be again.

In an important sense, therefore, the world today is predominantly a product of the historical path that flows through ancient Greece and Rome, thence through the Islamic World and China, and into medieval and modern Europe, especially as filtered through the Reformation and the Enlightenment. The Reformation occurred when Luther and his followers created the basis for Protestant Christianity beginning in 1517, and it is important because it broke the unitary intellectual framework and cultural domination of the Catholic

Church and thus enabled a cultural pluralism that might otherwise have been quite difficult. All things equal, cultural pluralism and competition tend to lead to more rapid social and cultural evolution and experimentation—and thus technological innovation, which can be easily stifled where specific worldviews gain political and cultural dominance. Interestingly from an engineering perspective, Protestantism, which focused on an individual relationship with God rather than one mediated by specialists (e.g., the Catholic clergy), could only thrive because of significant advances in the technology of printing, which in turn made the Bible widely accessible (a development that in its turn depended on the social development of widespread literacy, the "soft" side of printing technology). Lest it be thought that technological innovation is necessarily benign, it should be remembered that the following centuries of European history were among its most bloody as different Christian belief systems warred across the continent, aided by technological developments such as gunpowder, cannon, and the development of heavily armed ships.

Religious warfare between countries, and violent and bloody internal battle when the crown had different religious beliefs than members of a society, played a large part in the evolution of the most important institution of the European experience, which is still the dominant form of political organization and the only sovereignty recognized internationally: the **nation-state**. The hyphen tells you that a "nation" and a "state" are not the same: a "state" is a political entity such as a recognized country; a "nation" is a cultural or social entity such as an African tribe. A "nation-state" such as the United States, China, or France, exists when the two are coextensive. In some cases, such as in Africa where country borders may not reflect cultural or tribal borders, one may have a "state" without having a "nation"—a condition that frequently leads to civil war and continuing conflict.

The idea of the nation-state arose as a product of the Peace of Westphalia (more specifically, the treaties ending the Thirty Years' War signed in Osnabruck and Münster in 1648). It also is not necessarily an unmitigated blessing: The French Revolution and Napoleon created modern mass nationalism out of that framework, forces that the German state turned into a massive global war for the first half of the twentieth century (although World War I and World War II are separated in the history books, many military historians regard them as incidents in the same basic conflict). Nonetheless, nation-states are today still the only recognized sovereign entities in international affairs: To be a member of the United Nations, you must be a country, and Africa still struggles with the discontinuities that arose when the underlying tribal system that characterized that continent was divided somewhat arbitrarily into countries by European imperialists.

The culture that Europe globalized around the world can be summarized as being that of **the Enlightenment**. The Enlightenment, the period in

European history around the seventeenth and eighteenth centuries when reason, especially reason supported by scientific experimentation and observation, replaced religion as the source of authority and the primary mechanism for exploring human and natural systems, is also known as the Age of Reason. Of particular relevance to engineers, it includes as important values an emphasis on data, quantitative and rational analysis, and the scientific method; these values replaced reliance on written authority such as the Bible, and accepted historical scholars such as St. Thomas Aquinas, St. Augustine, and Aristotle. Before the Enlightenment, you proved you were right by referencing the appropriate passage in an ancient and accepted book; after the Enlightenment, you proved you were right by experiment and data. Engineering, although it existed before the Enlightenment (especially in military situations), is very much an Enlightenment activity. Additional values of the Enlightenment include rationality as a hallmark of cognition and an important value in itself, individual autonomy and individualism, and pluralism and democracy.

Equally important, the market-driven capitalist form of production that dominates world economic activity is a product of the West, as is the scientific and technological culture that accompanies it. Indeed, even the greatest challenge to market-based capitalism, Marxism, is a worldview that itself grew out of European experience. Marx viewed himself as an Enlightenment scientist, using scientific principles to explain how class struggle would inevitably lead to the evolution of a communist society, even as two hydrogen atoms and one oxygen atom connected together made water.

These observations should not be read as rejecting or devaluing the scientific and technological developments of other cultures, such as Islam during the early Middle Ages or the well-documented Chinese innovations in technology, nor of the contributions to Eurocentric culture made by many other societies, cultures, religions, and belief systems. Indeed, the Islamic World was clearly ahead of early medieval Europe in science and technology, and China was at least equal in many ways until the 1700s and even 1800s. But it was in Europe, especially the rationalistic and scientific Europe of the Enlightenment, where the threads of institutional evolution, religious conviction, scientific and technological capability, cultural systems, and economic theory and market development, all came together. The result was not just the development of ad hoc technologies or scientific discoveries, but the generation of a cultural juggernaut—a set of uniquely powerful networked ways of thinking, technologies, and institutions—which swept over the globe. If you are ever tempted as a sustainable engineer to underestimate the power of what you do, remember how and why we have come to sustainable engineering, and how complex the concept of technology really is.

BUT TWO CRITICAL OBSERVATIONS TEMPER THIS STORY. The first is that, as even superficial observation indicates, it is a profoundly multicultural world. In

part, this is because the Westernization of the world has not occurred explicitly in many cases, but at a more subtle level, in ways that appear to the observer to be culturally neutral. Examples might include the idea of progress, time as a linear structure, the technological and scientific underpinnings of modern life, the market economy, and the idea of individual human rights. Moreover, especially with the end of the Cold War, challenges to different aspects of this model have proliferated. Radical Islam, fundamentalist Christianity, and deep green environmentalism challenge the alleged benefits of market economies and modern technologies; some religious groups challenge equal rights for women, gays, and nonbelievers; many communities challenge the traditional nation-state model by demanding increasing devolution of authority. Although it can still be argued that global systems indeed reflect European Enlightenment roots, it does not appear that they are therefore more homogenizing; rather, there seems to be a dramatic increase in social and cultural complexity, with a patina of globalized cultural similarity. The European Enlightenment may be a global culture, but in many cases it is an addition to, not a substitution for, increasingly complex and assertive local cultures, both new and ancient. Why does this matter to a sustainable engineer? Quite simply, because engineering always in some way involves design, and design is a cultural activity.

A second observation is that the Enlightenment as global culture has succeeded, ironically, because it uniquely carries within it the potential for continued challenge and learning. Marx may have challenged from inside; the aesthetics of Africa and Asia challenged from outside (Western art—indeed, the career of Picasso alone—is a history of learning from other aesthetic traditions and incorporating them into new visions that then become, over time, accepted art, which is reinvented by incorporating yet again external cultural influences). Enlightenment culture, in part because it emphasized rationality, has proven very adept at adopting the best practices and ideas from around the world. It is not unique in this: Japan, once it had opened to the West, was extraordinarily competent in adopting successful technologies and cultural patterns. But that, taken together with its other characteristics, is an important, probably absolutely necessary, component of the success of Eurocentric Enlightenment global culture. Why does this matter to a sustainable engineer? Quite simply, a good engineer, just like the Portuguese building their caravel, will draw the best ideas she can from anywhere she can get them, challenging herself just like the Enlightenment challenged, and continues to challenge, itself.

2.2.1 Multiculturalism and Technological Evolution

The accelerating complexity of the world is an important challenge for sustainable engineering, or for that matter to any who choose to do engineering of any sort going forward. Natural disasters and epidemics, and their associated

economic and social effects, are more challenging given modern transportation infrastructure and globalized patterns of commerce and travel. Modern terrorism is not a new phenomenon, but terrorist access to the technologies of weapons of mass destruction, and conversion of familiar civilian technologies to terrorist ends, is. It used to be that those who designed weapons systems did so for the state; now they may be doing so for nongovernmental actors, whether called terrorists or freedom fighters. Cultural conflict is as old as historical records, but the Internet and ubiquity of exposure to others—a function of modern information and communication technology (ICT) networks—creates a condition where a few cartoons in a small northern European country, Denmark, can ignite global unrest. Note well that each of these phenomena has a powerful technological dimension and thus challenges engineering in new and unique ways; hence, then, the interest in sustainable engineering. Engineering by itself will not "solve" any of these challenges, whether natural disaster or terrorist incident, but it can reduce the probability of their occurrence, reduce the immediate cost of each incident and emergency response, and facilitate recovery. Each of these patterns could not exist without the enabling technology, from new cities built in coastal floodplains to the Internet to airplanes as major contributors to the spread of flu and other diseases. Engineers can no longer escape considering the uses to which their technology may be put in an increasingly complex and multicultural world.

Ironically, perhaps, these sorts of challenges have been enhanced since the Cold War ended, because the collapse of the global bipolar order that the Soviet Union and the United States imposed during that period has created a social and political space for hitherto suppressed cultural conflicts to explode. In some cases, such as some communities in China and India, the explosion has been in technological capability and economic productivity, at least in certain subsets of their populations (a process that, by creating greater inequality, generates internal tensions in such cultures). In other cases, it has been far less benign. The terrorist attacks in the United States, Russia, Spain, India, Indonesia, and elsewhere are probably the most familiar, but it is important to recognize that around the world religious, nationalistic, and ideological fundamentalism has gained strength. In part, this is a reaction to a rate of unprecedented social, political, economic, and technological change, a desire to find order in an environment where not just minor social patterns, but deeply held foundational beliefs, are increasingly under threat, frequently from technological innovation.

It is very important for the technologist to understand this last point and not be puzzled by the seemingly irrational prejudice that some have against technology. There is a reason for this reactionary response: Any suitably powerful technology has strong cultural and social implications. Technology is not just widgets; it is change, and choices about some futures that become more

likely, and some beliefs and practices that, partially because of technology, become obsolete. This is something that many engineers fail to grasp, but that those affected by technology certainly do.

A familiar example is the birth control pill, which as a technology did not just affect the immediate question of avoiding pregnancy, but did so in a way that changed the power relationships between the sexes, which in turn challenged some religious traditions that viewed women as inferior to men. What is the responsibility of the engineers and technologists who created that technology? This is not just of historical interest, nor simply a question limited to birth control technology. Clashes between conservative institutions and technology continue: the Bush administration in the United States severely limited stem cell research for religious reasons; some members of the European Union continue to try to stop all use of genetically modified organisms in agriculture; some environmental groups have called for "moratoriums" of unspecified length on research and innovation involving core technologies such as nanotechnology. The sustainable engineer will realize that such calls, even if understood as ineffective, unrealistic, and ideological, are nonetheless an unavoidable part of the context within which modern engineering occurs. Further, this engineer should not write off such critiques, but use them to try to understand, and moderate through good design, negative dimensions of otherwise desirable technologies. Bans may be unlikely, but one often learns much more from criticism than agreement. More to the point, it is the design and deployment of new technologies that, partially because of the accelerating rate of technological change overall, and partially because many of these technologies challenge embedded social and cultural beliefs, play a major role in generating such responses. Helping society manage such stresses, therefore, is an important part of the sustainable engineer's remit.

2.2.2 Changes in Governance

As previously mentioned, since the 1600s the nation-state has been considered the primary source of authority in global affairs. Increasingly, this viewpoint seems quaint and over-simplistic, both because cultural divides make fewer states coextensive with nations, and because many communities of interest are arising that cut across the jurisdictional level of the state. This is a fairly new phenomenon, reflecting both technological and cultural changes. Indeed, thirty years ago when the author of this book began practice as a corporate environmental attorney for an American firm, power did reside in the nation-state: When an environmental issue of any significance arose, one generally went to Washington, DC, or another country's corresponding capital, and worked with the federal authorities under whom that subject resided (U.S. Environmental Protection Agency, for example). Political and regulatory

compromises with environmental groups, affected communities, and others were worked out in arenas defined by the state, and formalized in federal regulations which flowed down the traditional government paths: to the provinces, to localities.

BUT NO MORE. As Figure 2.2 suggests, in many cases today important policy issues are worked out directly with interested stakeholders and communities, and those communities may well be virtual as well as geographically defined. Moreover, some national governments are shifting responsibilities and activities that were previously reserved for them over to either private firms or **nongovernmental organizations** (NGOs, such as environmental or human rights groups). The United States, for example, privatized warfare to an extraordinary degree in Iraq and Afghanistan, while the European Union now flows a great deal of its governmental foreign aid through NGOs. Space flight, heretofore reserved exclusively for governments, is increasingly the province of industrial interests. The beginning of the privatization of space occurred, somewhat ironically, as the Russians flew "space tourists" in their workhorse Soyuz spacecraft; in the United States a number of firms are planning and building private launch facilities. The respective governments regulate, but they are no longer exclusive operators. This does not mean that the state system and individual nation-states are becoming unimportant; rather, it is that

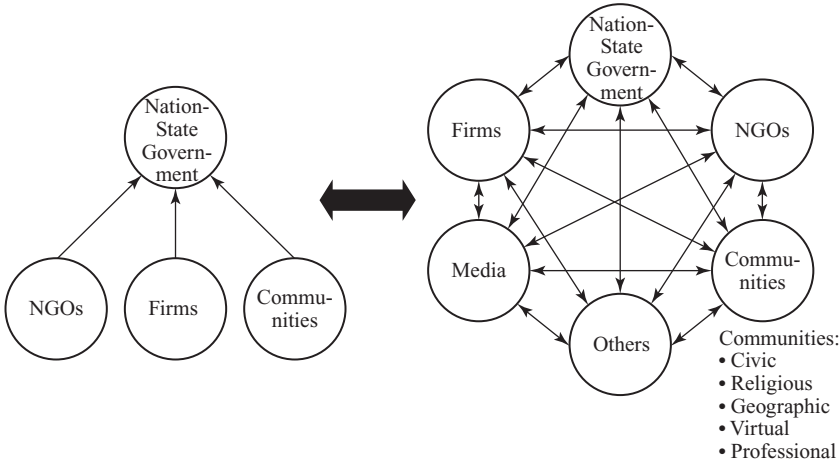


FIGURE 2.2 Shift in Governance Structure. This figure illustrates the important shift in governance structure from one focused on the nation-state as ultimate decision maker, to one where a number of different entities are independent centers of intersecting authority. Those shown, of course, are not exclusive; depending on the issue, one might include global criminal interests supported by legal restrictions on markets that enable monopoly rents, religious entities, nongovernmental military organizations, and the like.

the governance structure is becoming much more complex, so that states are no longer the only entities involved.

THE IMPLICATIONS FOR SUSTAINABLE ENGINEERS ARE OBVIOUS. Sustainable engineering requires an ability to understand which groups of stakeholders are important to a particular decision, to work with all of them, and to respect their positions, without imposing one's own particular worldview. This responsibility arises because the sustainable engineer must work in the real world, and that world is now complex not just technically, but culturally, socially, and politically. In addition, robust solutions require not just technological competence, but the ability to integrate the interests of many different individuals and organizations, even when they may disagree profoundly. As Figure 2.3 illustrates, there are a lot more NGOs out there than before, and they don't agree even among themselves. One example of this arises when wind or solar farms, often located far from demand centers, need to build new high power lines: Such a case in Los Angeles ended up with some environmentalists bitterly opposing any new power line construction, and others arguing in its favor, because the power to be generated by the proposed solar facility in the deserts east of Los Angeles would be carbon neutral. And this example was just among the environmental groups: Each affected community, each firm that might gain or lose by the contracts, each local and regional government, had its own position as well.

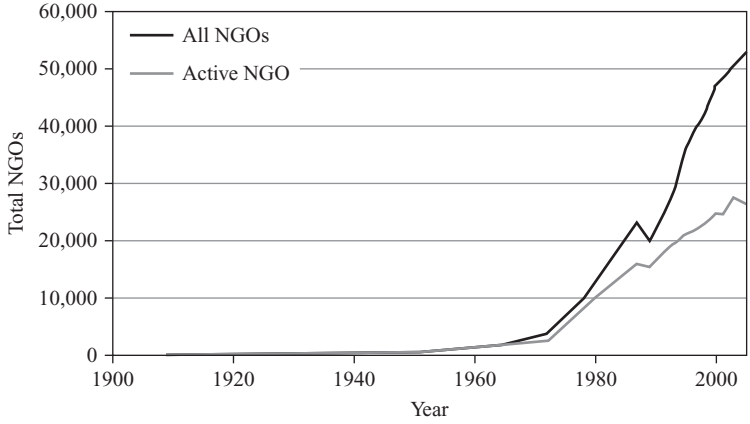


FIGURE 2.3 Number of Nongovernmental Organizations. This graph shows the growth of nongovernmental organizations, or NGOs. Such institutional growth implies proliferation of concerns and issues that must be dealt with as part of the management of large projects, a complexity that greatly complicates life for the engineers and managers involved in them.

(Data from: Union of International Associations. "Statistics Compiled by the UIA." http://www.uia.be/stats: accessed July 8, 2009.)

Moreover, the shift in governance does not just mean that more and different stakeholders are involved in engineering projects and issues of technology policy; it also means that expectations are changing. Most obviously for many engineers, the trend towards privatization, combined with the increasing power of emerging technologies, which are generally the province of private firms, is changing public expectations about the role of firms. It used to be the case that most people accepted that the firm was essentially a profit-oriented institution, and so long as it approached this mission responsibly and ethically, and was fair to its employees and customers, it had performed its social role. More recently, however, some cultures increasingly make the firm responsible for the systemic implications of any technology it participates in, as well as social responsibilities that may extend to supporting ends such as education and public health, and even "sustainability" generally. Some have noted that this represents jousting between the American ideal of the firm, and the continental European social democratic ideal of the firm. At any rate, it makes for a more complex environment, especially during the transition between different models of the firm. The case outlined in Box 2.1 gives one example, but there are many others. We are in a period not just of technological change, but of institutional and governance change; because the technologist is expected to provide leadership and advice to society in this very complicated and changing context, understanding these trends to the extent possible becomes a core competency of sustainable engineering.

2.3 INFORMATION AND COMMUNICATIONS TECHNOLOGY AND SYSTEMS

A key trend in the anthropogenic world, particularly in terms of its increasing complexity, is the continuing change in **information and communication technology**, or ICT. ICT includes computer, smart phone, consumer electronics, and transmission technologies; the Internet in its various guises; storage and entertainment media; software packages and services such as social networking services, network security packages, and search engines; and much more. As computation becomes more universal, the boundaries of ICT become somewhat difficult to define: a modern car, for example, is in many ways essentially a computer on wheels, and the "smart electric grid" is only the most obvious Internet-enabled infrastructure system. More exotically, the human genome project and feats such as designing and building bacteria from scratch suggest deep connections between biological systems and biotechnology, and ICT. Nonetheless, there are a few basic points that are best emphasized up front.

Most important, the amount of information in the world is expanding with extraordinary rapidity. The University of California at Berkeley 2003

BOX 2.1 Monsanto, Genetically Modified Organisms, and Corporate Responsibility

Monsanto, currently known as a producer of bioengineered seeds and agricultural chemicals, was initially founded in St. Louis, Missouri, in 1901, primarily as a supplier of raw materials to Coca-Cola. Since then, it has ranged broadly depending on market demand and opportunity, having at one time or another produced plastics, synthetic fibers, defoliants for the United States during the Vietnam conflict, and many agricultural products, such as dichlorodiphenyltrichloroethane (DDT). Beginning in the mid-1990s, the company underwent a series of mergers, acquisitions, and spin-offs to transition to a biotechnology firm.

It is this phase of Monsanto's activities that offers interesting insights into sustainable engineering. The political and cultural history of the GMO ("genetically modified organism," usually used in reference to agricultural products such as biotech corn, rice, soya, and other crops) controversy is extensive and fascinating, and far too complex to cover in a short piece. In brief, however, it began when Monsanto developed GMO corn that was resistant to certain pests, thereby reducing the need for pesticide application. This biotech corn was generally accepted by the American public and most regulators (including European regulators), but attacked by European environmentalists, especially Greenpeace. There is little question that the campaign against GMOs was highly effective, even if characterized by significant misstatement of fact (the "Frankenfood" anti-GMO campaign featured environmental activists dressing up as dangerous food, and brilliantly connected the archetypal image of "mad science" from the Frankenstein sagas to every consumer's diet. There is no evidence, however, that GMO foods are any more or less dangerous than any other kind of food; many scientists regard the real questions around GMOs as involving their impacts on ecosystems and nontarget organisms). The corporate response, by contrast, was highly ineffective, and too often sounded condescending, paternalistic, and defensive. This controversy continues today: The European Union and its environmental supporters continue to oppose GMOs and GMO technology, with the more aggressive European activists destroying experimental crops and research facilities. GMO crops, however, are widely planted in other countries, especially those countries where industrialized agriculture is the norm, such as the United States, Brazil, Canada, Argentina, and increasingly China and others. GMO adoption in terms of acreage planted has been strong and steady since the products were introduced in 1996 (see Figure B2.1.1).

Why Such Strong European Resistance?

There are a number of factors, including antipathy towards large American companies perceived as interested only in profits, a very effective marketing campaign by Greenpeace and others, public distrust of government food regulation arising from previous incidents such as "mad cow disease," and a strong sensitivity to any form of genetic manipulation deriving in part from the infamous practices of Nazi scientists during World War II.

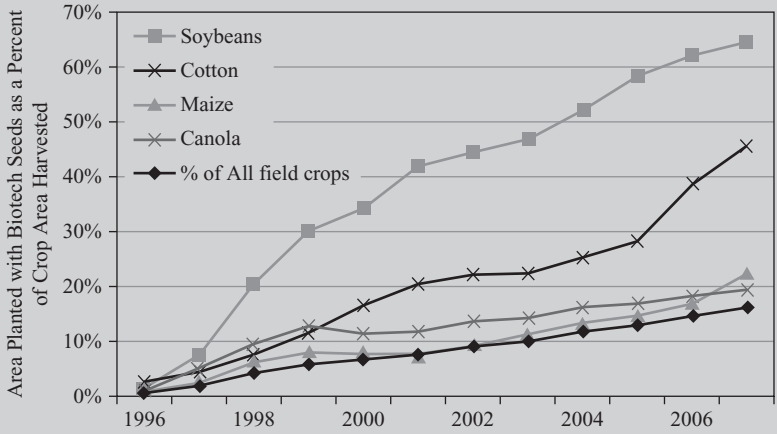


FIGURE B2.1.1 Global Area Planted in Biotech Crops. Here, the rapid adoption of genetically modified agricultural organisms is illustrated. These figures are even more impressive given European opposition to GMO technology.

(Data from: James, C. *Global Status of Commercialized Biotech/GM Crops: 2007*. Ithaca, NY: ISAAA, 2007; United States Department of Agriculture Foreign Agricultural Service. "Production, Supply and Distribution Online," USDA. http://www.fas.usda.gov/psdonline/psdQuery.aspx: accessed May 25, 2010.)

More important, perhaps, this case illustrates that different cultures have different ideas about firms and their responsibilities. Monsanto, reflecting its American origins, saw the firm as responsible to its shareholders, and thus as a profit-making entity. Many Europeans saw the firm as responsible for the broader social and environmental implications that the underlying technology might entail, and further argued, based on a concept known as the **Precautionary Principle**, that no new technology should be introduced unless it could be demonstrated to be less risky than the status quo (further discussion of the Precautionary Principle as an ethical framework is provided in the chapter on macroethics and sustainable engineering). Monsanto's misreading of this cultural difference was compounded by a response that was highly technocratic: The engineers and scientists knew the answers, and the public should not challenge their expertise. This hubris, and a refusal to dialog with those elements of the public that might have been open to discussion, ensured political difficulties.

Finally, many Europeans have a very different view of agriculture and food than Americans. For many Americans, food is a commodity, low price is a virtue, and thus the proper production model for food is industrial. For many Europeans, agriculture is not just about production, but about landscape, and an emotional and psychological component of their self-image and history. Thus, for example, the Swiss model of agriculture, with a few picturesque cows in an alpine environment, is not cost-effective, but it maintains a standard of living, a culture, and a landscape that are themselves important values (and not just cultural: tourism is an important sector of the Swiss economy). This difference in perspective leads Americans and Europeans to evaluate agricultural production technologies very

differently: Generalizing somewhat simplistically, Americans look for efficiency; Europeans look for cultural attributes.

Why Does This Matter to the Sustainable Engineer?

Quite simply, you don't want to be the one who unnecessarily contributes to the sort of conflict that GMOs continue to be caught in. It is usually unnecessary, it is not smart, it can damage your client or firm, and it can stifle the development of technologies that might be otherwise socially and environmentally valuable. You also want to be sensitive to what different cultures may expect from your work: in some cases the differences between, say, Europe and America loom large; in others, not so much. And you need to be sensitive to stakeholders. Although it is doubtful that Monsanto and Greenpeace would have had any common ground, there were many other organizations involved, and some dialog was certainly possible. Finally, be sensitive to changing expectations, especially about the role of the firm. It used to be generally accepted that the principle objective of the firm was making money responsibly and ethically, but that is changing, especially where new technologies are involved.

study, "How Much Information?," which tried to quantify the Information Revolution, estimated that the rate of "new" information stored was growing at about 30% a year between 1999 and 2002, implying a massive increase in available information. Whether information is "new" or simply recopied is, of course, not only difficult to determine in many cases, but of interesting relevance: If a song is recorded in a location where it is freely accessible as an electronic file, as opposed to located on a CD where one must purchase the entire CD at highly inflated rates to hear it, the implications for the existing recording industry may be significant, even if copies of the original are not "new" information at all.

What is clear is that significant shifts in information patterns characterize our modern age. Figure 2.4, for example, compares the amount of information transmitted by traditional mail, as opposed to that transmitted by email. The dramatic increase in information transmitted per capita is clear; the social and cultural changes implied by this increase are still unclear, although they may well be quite important over time. Moreover, given that communication increasingly includes information-intensive modalities such as video, this trend may be expected to continue (at least directionally). The surge in information generation and transmission is also much greater than this figure implies because of the rapid rise of social networking services.

Despite the many new information and communication technologies, services, and sources that are rapidly being commercialized, however, much information continues to be provided through traditional channels such

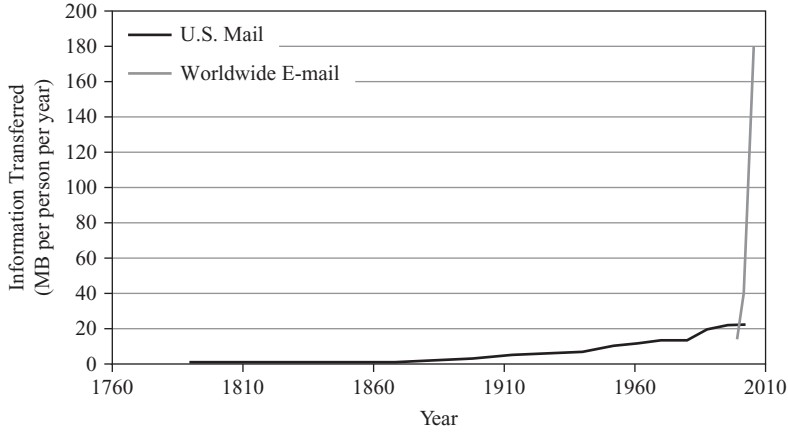


FIGURE 2.4 Written Communication. This graph assumes that the average piece of U.S. mail contains two pages of written content, each containing 15 kilobytes of information per page; it further assumes that the average email message contains 59 kilobytes of data. All data includes marketing material and spam. Note that data for mail is from the United States only whereas email is worldwide. Although this is not ideal, both data sets have been normalized on a per capita basis and should therefore be comparable. The growth in email information flows per capita is even more remarkable because it is an average that includes many regions of the world where email is not readily available.

(Data from: Lyman, P., et al. "How Much Information?" 2003. http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/: accessed July 9, 2009; United States Census Bureau. "Total World Population," 2008. http://www.census.gov/compendia/statab/cats/international_statistics/population_households.html; accessed July 9, 2009.)

as television and radio (see Figure 2.5). Nonetheless, even here technology shifts are significant; cable offers many more channels than over the air television technology, and satellite radio similarly expands the options available to consumers.

This information explosion is fed by rapidly accelerating development of technological capability and new platforms (see Chapter 8), which are themselves quite impressive. What is important to understand at this point is that the sheer volume of information itself is having substantial cultural impacts, because culture is a domain heavily dependent upon and reflective of information structure. The more information that people have to process, the more they are likely to need ideologies and worldviews that help them simplify the world so they can perceive it and act in it relatively ethically and rationally. Such a dynamic raises at least the possibility that an information-dense world is also one where levels of conflict are higher, because different worldviews are more likely to proliferate. In addition, the more information generated and gathered around each person, from credit card expenditure data, to Facebook and social networking data, to pictures and recordings from public surveillance cameras and random individuals all able to take cell phone pictures and

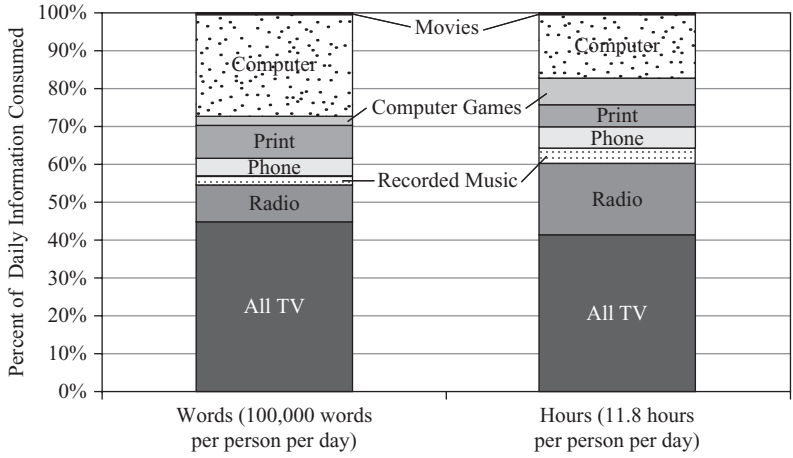


FIGURE 2.5 Information Consumption in the United States by Source in 2008. This chart shows the time and quantity of information that Americans receive from various sources. Note that some information streams may be consumed simultaneously. For example, if someone has the TV on while working at a computer, that information would be double-counted in this figure.

(Data from: Bohn, R. E. and J. E. Short. "How Much Information? 2009: Report on American Consumers," 2010. http://hmi.ucsd.edu/pdf/HMI_2009_ConsumerReport_Dec9_2009.pdf: accessed October 20, 2010.)

video, the more problematic any concept of privacy becomes. But the sustainable engineer should not make any assumptions about universal cultural norms regarding issues such as privacy: Data indicate that older cohorts care more about privacy than younger "digital native" groups do, and some cultures, such as Europe, are more inclined to use government to protect privacy values than others.

2.4 TECHNOLOGY AND SUSTAINABILITY

From the perspective of sustainable engineering, it is important to understand that technology and technological evolution have a critical relationship to any sort of sustainability, although issues of technology are often lacking in sustainability discussions. This relationship can perhaps be best understood using the well-known and traditional **Master Equation** illustrated in Figure 2.6, which qualitatively breaks down human environmental impact into three components: population level, consumption per unit population, and impact per unit consumption.

Even in this simple form, the Master Equation is useful, because it demonstrates conceptually that environmental impact does not stem only from population growth, or from consumption, but from the way the two are related. Much more important from the perspective of sustainable engineering,

$$\text{Overall Environmental Impact} = \text{population} \times \frac{\text{resource use}}{\text{person}} \times \frac{\text{environmental impact}}{\text{unit of resource use}}$$

FIGURE 2.6 The Environmental "Master Equation." The familiar Master Equation qualitatively breaks environmental impact down into three factors: population, resource use per unit population, and environmental impact per unit of resource use. In this form, this is a highly idealized treatment of environmental impact, but still useful for understanding different pressures on environmental systems and how they are related. This equation is commonly associated with the field of industrial ecology, and is sometimes called the IPAT equation, because it connects Impact, Population, Affluence, and Technology.

it identifies the key role of technology and thus engineering in the process. Population growth, although perhaps impacted by technology (e.g., birth control mechanisms), is clearly not a variable subject to rigorous control through technology; moreover, this term is still growing strongly (that is, global population continues to increase). The second term is basically a consumption term, and consumption itself tends to be culturally determined. It also is not subject to control directly through engineering, and it is similarly growing strongly globally. It is the third term, then, that is a technology term, because the environmental impact of consumption is determined to a large degree by the design of the technologies—the engineered systems—that people have available to meet their needs. Moreover, the equation taken as a whole strongly suggests that, especially in the short-term and given continued robust growth in terms one and two, better engineering and design are critical to controlling environmental impact.

But the alert reader will have already noted that the Master Equation in this form is inadequate from a sustainable engineering perspective for at least two important reasons. First, it speaks only in terms of environmental impact, not in terms of the integrated human/natural/built systems that are the essence of sustainability in the Anthropocene. Sustainability is much larger than just environmental issues; it also includes social, cultural, institutional, economic, and technological dimensions. These dimensions are not included directly in the Master Equation. A second limitation of the Master Equation for purposes of sustainability arises because it is framed in terms of "resource use." Resource use—material and energy flows—are indeed the traditional focus of environmental policy, but this perspective is inadequate in a world where information systems and emerging technologies are redefining so many human activities, including, to some extent, the human itself (see Chapter 10). Resource use is important, but unless services, information systems, and the social and cultural dimensions of earth systems are taken into account, we are avoiding the true challenges of our time by focusing on relatively trivial parts of a complex system. The Master Equation is useful as far as

$$\text{Sustainability Impact} = \text{population} \times \frac{\text{quality of life}}{\text{person}} \times \frac{\text{sustainability load}}{\text{quality of life unit}}$$

FIGURE 2.7 The Sustainable Engineering "Master Equation." Note that this reconceptualized equation is not, as in Figure 2.6, concerned only with environmental impact, just as sustainable engineering is more than environmental engineering. Note that further differentiation of the third term—by, for example, breaking "sustainability load" into quantifiable elements plus a residual term—would enable the sustainable engineer to identify specific targets for improvement.

it goes, but it does not go nearly far enough; for purposes of sustainable engineering, it is too limited.

Accordingly, the Master Equation must be updated; our "Sustainable Engineering Master Equation" is presented in Figure 2.7.

Figure 2.7 confirms the importance of technology and sustainable engineering, in that the third term remains a critical technology term where good design and good engineering can reduce loading on sensitive natural, social, and cultural systems. But this equation does more than that. By emphasizing quality of life, which for most people today includes environmental quality, it does not privilege one value—say, the environment—over all others—say, jobs or individual freedom. Moreover, quality of life includes all aspects of both economic and noneconomic activity; it is not just a function of resource use, but of the many intangibles that affect individuals and their happiness. It also allows for the sorts of tradeoffs that sustainability advocates encourage, such as a high quality of life with less material consumption, but more leisure time for appreciation of culture and arts. This breadth and perspective, however, also explains why sustainable engineering is more challenging than environmental or "green" engineering, because although it is relatively easy to quantify and consider many environmental variables, such as emissions or energy consumption, it is far harder to create good designs able to meet the multiple demands of society, culture, environmental quality, and economic and technological performance.

2.5 DEMOGRAPHIC TRENDS

The last chapter discussed the dramatic increase in overall human population, especially since the Industrial Revolution. Such growth, of course, contains a number of interesting subthemes, two of which are particularly important to sustainable engineering: urbanization and changes in age structure. Without intending to slight other issues that require attention, such as differences in distribution of jobs and wealth within and among societies, a brief discussion of these two areas may serve to sensitize the sustainable engineer to the importance of demographic changes.

2.5.1 Urbanization

The Greek philosopher Aristotle famously observed that "man is by nature a political animal," and reflecting the structure of Greek society of his time, that the essence of the human is the city-state. He added that people congregated in cities to live, but remained in cities to live the good life. Centuries later in the late medieval period of Europe, it became a common observation that "stadtluft macht frei"—"city air makes you free." Although this statement was true given the status of the serf in feudal society, it was also the case that pestilent city air, and the close quarters and primitive sanitation of medieval cities, as often killed a person. Compared to the countryside, however, at least that person died free, if younger. More generally, there is no artifact that is more characteristic of humans, and that is more intertwined with the evolution of the species as it now appears, than the city. Cities are the engine rooms of technological, economic, and cultural evolution. Any attempt to understand the anthropogenic world that does not include urban areas is bound to fail—and any engineer who doesn't understand that humans are, after millennia, finally becoming predominantly urban beings doesn't understand his or her world. Cities are, after all, complex highly engineered and technological environments; combine human and natural systems at all scales; are cultural and economic centers; have enormous implications as centers of material and energy flows; are nodes of significant consumption; are centers of regional and global ICT networks; and are where a large and growing human population is increasingly concentrated. This trend has profound implication for engineering, not just because of what it implies in terms of infrastructure needs, but in terms of resource management (including energy and water), information and communication patterns, management of pollution, maintenance of desired ecosystem state, and many other dimensions. This is especially true given the aging of the world's population, discussed below.

It is estimated that about eighty million people a year are moving into cities, mostly in the developing world—that's the equivalent of establishing eight new megacities a year for the foreseeable future (a megacity is, by convention, a city with a population over ten million). Figure 2.8 gives some idea of the staggering growth in cities with a population of over a million; most of these cities will be established in less-developed regions, and thus will imply urbanization with quite limited resources. The implications of this growth for engineering of all types—civil, transportation, information, consumer goods, automotive, and others—are obviously profound but, at this point, little understood and less taught.

Indeed, the United Nations estimates that the urban populations of Africa, Latin America, and Asia will double over the next thirty years, growing from 1.9 billion in 2000 to some 3.9 billion in 2030, at which point some 60 percent of the world's population will live in cities, a percentage that will

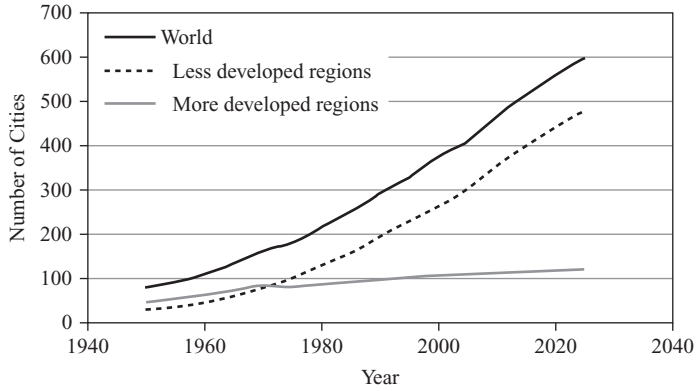


FIGURE 2.8 Number of Cities Having a Population of One Million or More. This chart shows the number of cities with a population over a million that have grown up since 1950, and will develop in the next decade or so. The implications of this growth for energy, water, and resource consumption are significant, but the industrial ecology of urban systems is an understudied area.

(Data from: United Nations Population Division. "World Urbanization Prospects: The 2007 Revision" [Document No. ESA/P/WP/205], February 2008: accessed http://www.un.org/esa/population/publications/wup2007/2007WUP_Highlights_web.pdf.)

continue to grow. As Figure 2.9 shows, the urbanization of developing countries will lag that of developed countries in percent (but not absolute) terms for the foreseeable future (see Figure 2.10). All things considered, it is this migration to urban areas in less developed regions that poses the greatest political, social, economic, and engineering challenges.

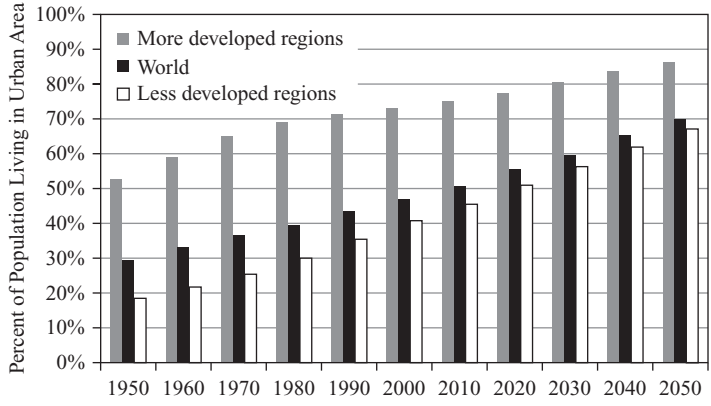


FIGURE 2.9 Urban Population by Development Region. Urbanization is occurring first in developed countries, a trend that was also true historically (industrialization in the United Kingdom coincided with a significant migration of population from rural to urban during the early Industrial Revolution).

(Data from: United Nations Population Division. "*World Urbanization Prospects: The 2007 Revision Population Database*", n.d. http://esa.un.org /unup/index.asp: accessed August 6, 2009.)

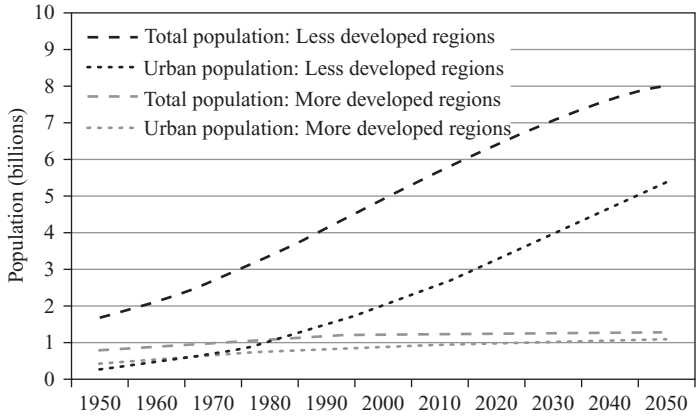


FIGURE 2.10 World Population by Development Region. Although the developed world will be more urbanized in percentage terms (Figure 2.9), the numbers of people migrating to urban areas will be far greater in less-developed regions. Because such regions generally have fewer resources, less resiliency, less employment opportunities, weaker social networks, and poorer physical infrastructure, the challenges posed by urbanization in less-developed regions will in most cases be far more substantial than in developed regions.

(Data from: United Nations Population Division. "World Population Prospects: The 2008 Revision Population Database." http://esa.un.org/unpp/index.asp?panel=2: accessed August 17, 2009.)

From a sustainable engineering viewpoint, note an important lesson about using and understanding statistics illustrated by this simple story. If one is interested in, for example, determining where urban populations are more dominant culturally, socially, and economically, one might choose the percent of overall population living in urban areas, rather than absolute numbers. If, however, one is interested in understanding where the greatest challenges of urbanization will arise, using figures based on the *percent* of population urbanized significantly understates the importance of developing regions. Rather, one would want to look at the *absolute numbers* of people migrating to urban centers, which is higher by several billion for developing regions. Neither figure is "wrong," but the choice of how to present the data makes all the difference in the world. Using data with integrity—an absolute requirement for sustainable engineering—is not just a matter of making sure the numbers are accurate; it also requires an understanding of what type of number presents the story in the most meaningful way (it also suggests that a sustainable engineer, who will often be working in highly conflictual situations, will want to ensure that the data presented are appropriate for the purpose at hand, rather than intended to mislead the analysis).

2.5.2 Urbanization and Sustainable Engineering

The difficulties of megacity governance, especially in developing countries, are well documented. The complexities of major projects within megacities, such

as the Boston Central Artery/Tunnel construction project, are also well known. These complexities arise in part from what might be called "traditional complexities" such as legacy networks and systems that must be accommodated as additional functionality is designed and built (a **legacy system** is one that is already in place; such systems operate as constraints on new designs and additions, regardless of the particular type of infrastructure).

But increasingly the complexity of cities arises from fundamental changes in what infrastructure is—in particular, from accelerating information complexity. Moreover, this evolution provides some important indications as to what sustainable engineering in the current era will have to be, and how it differs from the more traditional, disciplinary forms with which we are all familiar.

To begin with, we can learn from the general mischaracterization of urban systems by the environmental and sustainability communities. The historian William Cronon, among others, has noted that, for many environmentalists, the city has traditionally been viewed as the ultimate human environment, and therefore evil, in that it is the opposite of wilderness and "nature," which are understood to be the ultimate good. This bias has transitioned into much of the sustainability literature, leading to a widespread impression that cities are clearly "unsustainable," an interesting conclusion given that cities are perhaps the most sustainable artifact that humans have developed in terms of how long they tend to last. The failure to understand cities as evolving complex, adaptive systems, rather than static processors of materials and energy and producers of waste, is problematic in that it limits the usefulness of the dialog between sustainability advocates and urban systems just at a time when urban systems are about to experience a major leap in complexity. This will occur as three separate trends that, taken together, profoundly redefine the nature of urban systems.

The first is that cities, like virtually every other complicated human technology, are increasingly incorporating quite sophisticated information systems—including sensors, information processors and filters, networked computational capabilities, and the like—at a number of different scales. Increasingly, we are designing and building materials, components and subassemblies, buildings, city regions, infrastructures, and other urban subsystems to be "smart," with significant new information content and capability. Moreover, the information networks of these smart components, be they part of a bridge, a complete building, or even major infrastructure components such as water treatment plants, highway systems, or energy production facility, are not just internal but also integrate into larger information networks that may cover entire metropolitan areas. As well, every major urban system is increasingly an active node in complex regional and global information, transportation, energy, and financial networks, creating a level of information density, interconnectivity, and complexity that is entirely unprecedented.

At the same time, the structure and dynamics of ICT systems are changing fundamentally. Rather than the relatively passive, deterministic, and human-controlled technology that characterized the past (say, a telegraph or historic telephone network), new ICT systems are being designed to be self-defining, self-managing, self-configuring, self-healing, self-optimizing, and self-protecting, increasingly in real time. This so-called autonomic architecture and capability is currently being deployed by ICT firms at all scales, from chips, to computers to computer networks, to communication networks, to the Internet taken as a whole. Such capabilities mark a major leap in device and network functionality. Perhaps more important, however, learning capabilities are being built directly into these systems. For example, an intelligent network will "learn" its own characteristics, and thus be able to identify perturbations beyond normal operating fluctuations and respond in order to maintain operational functionality (e.g., respond to potential viruses as an immune system would rather than waiting for human intervention).

The third trend arises from increasing use and reliance on ICT systems across urban systems and involves the compromise of information systems, especially those that perform critical coordination and operating functions for important infrastructure, by both cybercriminals for blackmail, and national entities for political or military purposes. Tools can range from the unsophisticated and blunt, such as releasing Trojans or viruses or using "distributed denial of service" (DDOS) attacks to bring down a particular Web resource (a tactic used by Russian elements against Estonia in a now famous 2007 attack). Such attacks can range from fairly crude to very sophisticated; in some cases, state actors and their "privateer" quasi-state hacker corps insert "backdoors" into software and hardware systems. Such a "backdoor" is a set of instructions that allow the programmer or chip designer to manipulate or shut down functionality at some desired time in the future. The potential power of such backdoors was illustrated (perhaps, in the murky world of security and classified information) when Israeli jets bombed a suspected nuclear installation in Syria in September 2007, even though those installations were guarded by state-of-the-art radar facilities. But that radar failed just in time to enable the strike, much to the surprise of the Syrians. The hypothesized process is simple: The Israelis had purchased the code to instruct the Syrian radar to shut down. Given that much of the infrastructure currently being built in urban systems, including important components of the smart grid, uses equipment and products that are produced in many countries by many hands, the potential for introducing serious vulnerabilities into urban systems through inadequate cybersecurity designs is quite real. This is particular challenging because except for specialized computer engineers, very few engineers of any kind are given any training in information security design and implementation. From a sustainable engineering perspective, this is unacceptable; a design that works fairly well at the cost of introducing information system vulnerabilities into critical products or infrastructure can-

not be considered good engineering anymore. Note also that designing to avoid cybersecurity challenges is not limited to any particular nation: All nations currently jostling for geopolitical power—including China, Russia, India, the United Sates, Brazil, and many others—as well as many cybercriminals, are increasingly active across the Net, and thus across most designed systems today.

Consider now how these trends intersect and integrate in urban systems. Not only does the city become much more complex in terms of information density and networks, but so do the ICT structures supporting this complexity—exponentially so. The predictable result is not just an immensely more competent and functional city, but increased vulnerabilities, some understood; some just beginning to be understood, such as the threat of cybercriminals and cyberwarfare; and some not understood at all. For example, it is quite likely that such greatly enhanced ICT capability at all levels of the urban system will at some point begin to generate emergent behaviors and characteristics that are both unpredictable and potentially quite powerful, and ones that occur without passing through human institutions or filters. So, the sustainable engineer should begin to think about what the effect of all the systems increasingly being installed across urban centers will be as they begin to interact with each other—because at this point no one else is. Box 2.2 summarizes some of the implications of the emerging autonomic city.

Another instructive example of where urban engineering may go once it becomes more sensitive to overall system behavior is provided by the deceptively simple case of bicycles in urban environments, which shows how good engineering—sustainable engineering—is capable of addressing multiple social goals simultaneously, and doing so without violating principles of individual autonomy and choice. This example is suggested by even a casual comparison of practices in The Netherlands, a small, prosperous, and nicely designed society, and the United States, a much more autocentric culture. In The Netherlands, two things in particular stand out: the continuing popularity of bicycles, bolstered by an elaborate and effective bicycle transportation infrastructure, and a concomitant lack of obesity (to which one becomes all too inured in the United States).

Bicycles and a degree of health are not just related in the obvious way. Both can also be understood as emergent characteristics of engineering design and related policy choices, particularly involving transportation infrastructure. To oversimplify: In The Netherlands, bicycle infrastructure is given equal standing with automobile infrastructure, resulting in integrated transportation systems that as a spin-off, produce both environmental and health benefits. In the United States, we increasingly design infrastructure that actively discourages both walking and bicycle use, either directly (try walking from long-term parking lots to the terminal in many U.S. airports), or indirectly (you take your life in your hands if you bicycle along busy streets in many U.S. towns and cities). The United States is an obese society in part because of its design choices.

BOX 2.2 Lessons of the Autonomic City for Sustainable Engineering

For the sustainable engineer, this autonomic city case study illustrates several important principles. To begin with, it is clear that, as autonomous information system structures becomes ubiquitous at all scales, and as urban systems continue to become increasingly information dense within and among different levels, it will be impossible to determine what the city "knows" or how it will choose to react to changing conditions or threat. From this flows the principle that sustainable engineers must be much more flexible and adept at including uncertainty and unpredictability in the ambit of engineered systems in the urban context than in the past.

Second is the need to be able to work with and integrate across systems of systems, many of which will arise from different disciplines. Thus, for example, a recent test by intelligence officials involved hacking into an energy infrastructure information system and causing a major energy facility to physically blow itself up. Building responsible energy (and other) infrastructure therefore now requires not just mechanical and civil engineering competency, but ICT and software competency. Project managers must not only understand traditional metrics, such as efficiency of energy production and distribution, but new ones such as information system security.

More broadly, the enhanced ICT functionality of the city augments, rather than simplifies, the cultural, political, and social complexities of cities. Thus, modern urban systems are excellent illustrations of how, to be successful, engineering and management in complex systems become an ongoing process of dialog and learning, not a bounded problem-solving exercise. The system of concern is not static: You do not "engineer" a city and walk away; rather, you engage in a continual dialog with a complex evolutionary system that integrates human and natural components at many temporal, geographic, and informational scales. A city is clearly composed of human designs, both social and engineered; just as clearly, however, its evolution through time is neither predictable nor simple. Equally important, urban systems provide many cautionary examples of technocratic hubris, where we think we know more than we do. We think, for example, that we understand the information system that enables us to efficiently manage our energy infrastructure, but we have not looked at it from the perspective of a terrorist seeking to destroy it (or, according to intelligence agencies, organized criminal gangs that simply, and increasingly successfully, seek blackmail money). It thus cannot be overemphasized that the complexity, uncertainty, unpredictability, and contingent nature of the kinds of systems that engineers will increasingly be engaged in preclude simplistic, ideological, single-discipline approaches.

It is not that the sustainable engineer needs to become professionally competent at many different engineering disciplines: For many of us, one is about all we can manage. Rather, this case study suggests that knowing how to create profoundly multidisciplinary teams, work with other professionals closely and productively, and identify challenges, opportunities, and vulnerabilities emerging from the dynamics of the system as a whole, are critical skills for the sustainable engineer.

One shouldn't romanticize, however, about what's going on. The Dutch like their cars, and they don't eschew automobiles where appropriate, or require people to bicycle, or ban fast food outlets. Their approach is not to force people to give up what they think they want. Instead, in the best tradition of the free market and a trading society, they create an environment by building, for example, appropriate infrastructure, which offers a choice (unlike in the United States). In this way, they can increase rather than restrict choice, and thus create a higher quality of life while encouraging other public benefits such as good health (like most jurisdictions, tax and user fees are structured to encourage appropriate behavior). The key is in how engineering is framed: Is one designing roads, or is one designing integrated transportation systems, including public, private, bicycle, and walking environments? Much of sustainable engineering is framing the issues properly in the first place.

More broadly, once we realize that an increasingly important and appropriate goal of engineering is creation of option spaces that allow individuals to enhance their quality of life through choice rather than through mandate, other possibilities open up. For example, transportation engineers often talk of designing their systems for **activity fulfillment**, enabling people to perform the activities they want in as efficient a way as possible. Although in traditional transport engineering activity fulfillment is a way to compare different transportation infrastructure designs, this concept can be generalized to include "virtual activity fulfillment" through, for example, substitution of ICT options for physical transportation. Thus, young people increasingly substitute social networking on various Internet sites for driving to physical spaces such as malls to meet each other, while their working peers substitute virtual offices and network functionality for the traditional rush hour automotive jousting session. Increasingly, university classes and learned symposia, board meetings and engineering sessions, or art expositions and innovative concerts are now held in virtual realities rather than requiring cross-country or even international travel.

From a sustainable engineering viewpoint, what is significant in all these cases, from Dutch bicycles to Second Life seminars, is that the integrated natural/built/human systems that increasingly characterize our world can be designed in such a way as to provide additional choice to each individual while meeting overarching social goals. Provision of choice ensures both higher quality of life and a continuing respect for the individual, concepts that activists occasionally undervalue given their focus on specific issues. Additional choice may not have the same absolute guarantee of performance as mandates do—most obese Americans will not jump on bikes just because they are given a safe infrastructure, at least not at the beginning—but it builds public support for environmental and public health goals in a way that more adversarial and authoritarian approaches cannot. Gains in the short run may

be less under a choice regime, but they will be stable and enduring, and lay the groundwork for more rapid and fundamental cultural change than do adversarial and scare tactics.

But first, sustainable engineers have to build those option spaces. If we continue to think of built environments and systems as we traditionally have—transportation as roads, rails, and air, for example—we unintentionally restrict a panoply of benefits. Think of design as enabling choice and reflecting function, however, and we have the opportunity to enhance both quality of life and address a whole set of other policy dimensions. This case strongly suggests that integrated planning across transportation infrastructures, built environments that encourage safe walking and bicycle alternatives, and ICT networks and services that facilitate substitution of information for unnecessary transportation offers significant public benefits—not just economic and engineering efficiency, but public health and additional choice benefits. In a broader sense, however, this example provides one illustration of what sustainable engineering is, in fact, all about.

2.5.3 Age-Related Demographic Issues

Among the many age-related concepts that are of interest to the sustainable engineer, three stand out. The first two, the age structure of different nations and regions of the world, and the overall trend towards longer life, are important because they suggest different markets, different cultures, and different design choices. For example, if one is designing complex products that require sophisticated manufacturing techniques to be built in a developing nation with many relatively uneducated young people, one might want to substitute capital for labor, because the sophistication will need to be built into the manufacturing machines rather than assumed available in the working population. More fundamentally, of course, we might question whether such technology was appropriate at all for a country with plentiful labor but few capital resources (and probably little experience in keeping sophisticated technologies operating over time).

Many people don't appreciate how much age structures have changed over time. Let's begin by looking at the demographic transition that has occurred within one region, Europe. As shown in Figures 2.11, 2.12, and 2.13, the age distribution in Europe has shifted significantly over the past 55 years, and promises to do more so in the future. In 1950, Europe had a population distribution that implied a fairly stable overall population level: fairly even distribution across age cohorts until about 45, when increasing mortality began to take hold (Figure 2.11). In 2005, however, this had changed dramatically: There were less young people than middle-aged people, clearly signaling future problems unless work and financial institutions changed dramatically (for

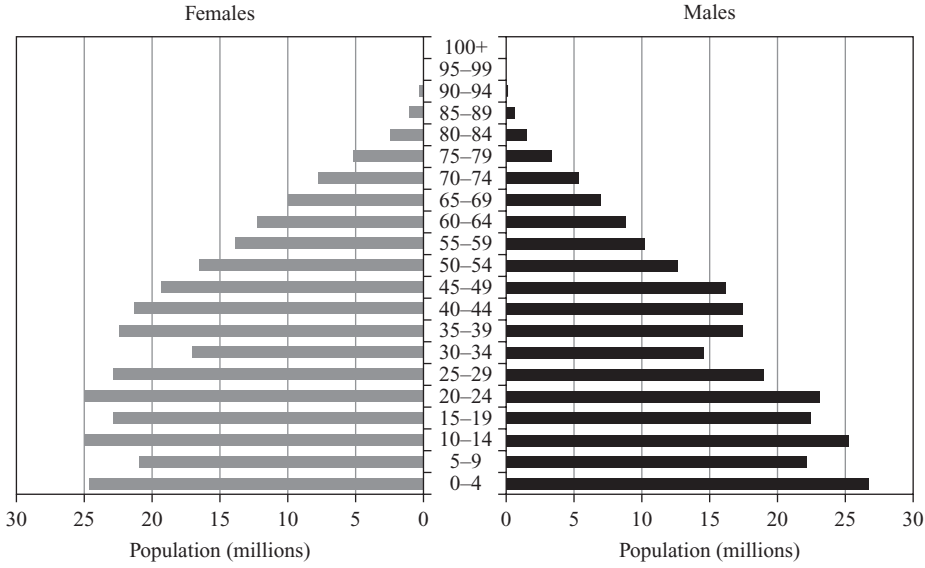


FIGURE 2.11 Population Age Distribution of Europe in 1950. Note that, despite World War II's impact on the 30–34 cohort, the shape is still dominated by the number of people under age 40. Compare this with Figures 2.12 and 2.13.

(Data from: United Nations Population Division. "World Population Prospects: The 2008 Revision Population Database." http://esa.un.org/unpp/index.asp?panel=2: accessed August 17, 2009.)

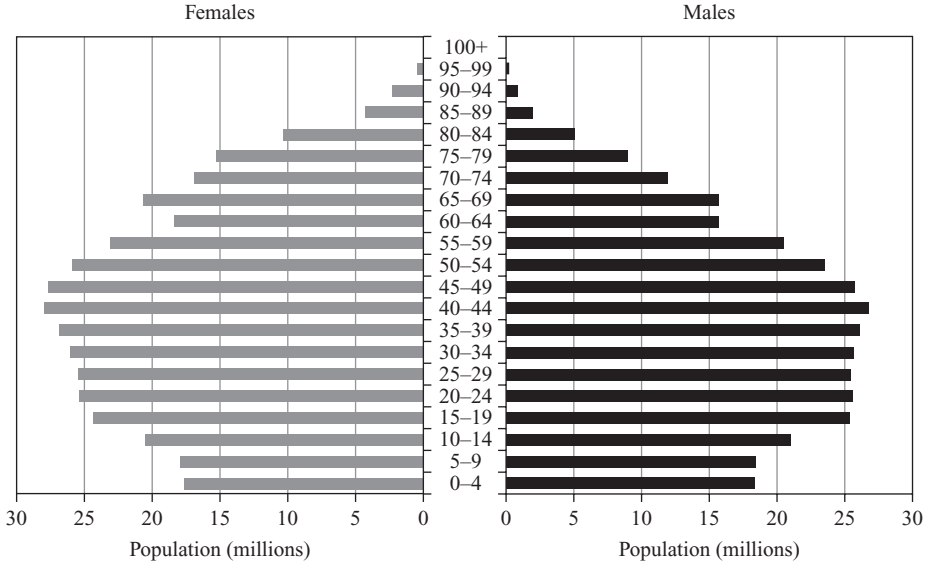


FIGURE 2.12 Population Age Distribution of Europe in 2005. Note that the youngest cohorts are far smaller than the bulge in the middle-age groups; this not only indicates an imbalance in European demographics from the perspective of employment management (when that bulge retires it will create serious economic issues because of the politics around later retirement and pension funding in Europe), but is in stark contrast to societies such as Egypt (Figure 2.15) and Africa (Figure 2.16).

(Data from: United Nations Population Division. "World Population Prospects: The 2008 Revision Population Database." http://esa.un.org/unpp/index.asp?panel=2: accessed August 17, 2009.)

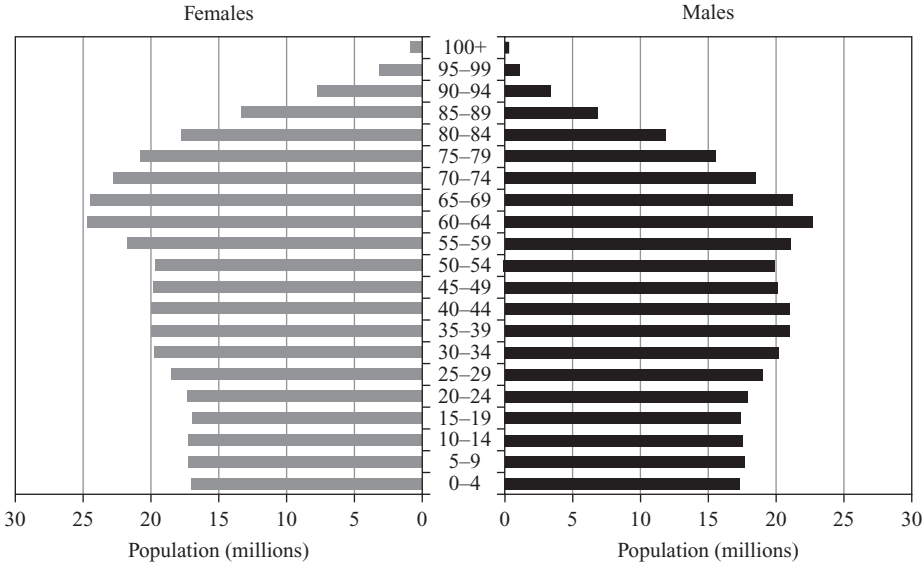


FIGURE 2.13 Population Age Distribution of Europe in 2050. The estimated age distribution for Europe in 2050 is highly unstable socially unless substantial changes are made in policy. In particular, it suggests that pension and employment practices will need to be changed significantly—put simply, people have to work longer, and/or current highly restrictive immigration controls may need to be loosened, so that immigration can balance low numbers of young Europeans. An alternative being explored by some other "aging societies" with similar demographics, such as Japan, is replacement of younger workers with robots.

(Data from: United Nations Population Division. "World Population Prospects: The 2008 Revision Population Database." http://esa.un.org/unpp/index.asp?panel=2: accessed August 17, 2009.)

example, adjusting to the reality that in the future fewer workers would likely be available) (Figure 2.12). But projections of future demographic structure are even more challenging. Figure 2.13 illustrates a population where the bulge is now in the 60 and beyond category; if all these people are retired, one could anticipate a fair amount of social stress from this figure alone. How many pensions for retirees, for example, is one worker able to support? Or, given the need for substitution of capital investment for labor implied by a shrinking work force, how much capital must one worker be able to access if he or she is expected to be supporting such a population structure? Or, alternatively, how long must older people be expected to continue working to avoid either personal or social financial collapse? Remember when considering such scenarios, however, that future projections are never reality: They are always based on assumptions that may or may not remain valid. For example, obesity might dramatically increase mortality among older cohorts; alternatively, mass migration from younger societies might increase labor availability by 2050. Scenarios such as Figure 2.12 are useful for encouraging creative thinking

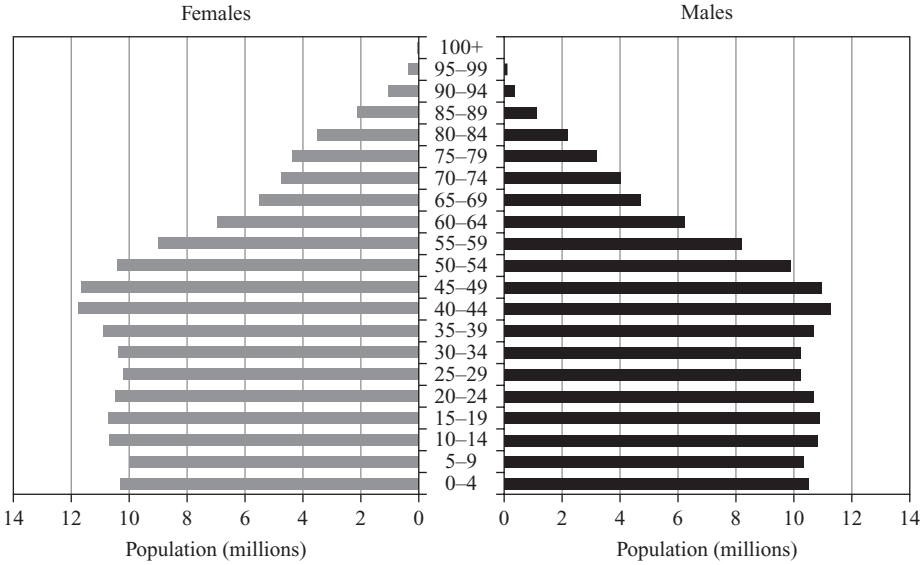


FIGURE 2.14 Population Age Distribution of the United States in 2005. As with Europe, the data suggest that pension, retirement, and medical insurance policies will need substantial change to respond to changing demographics; because these policies are so important to so many people, any such change will not be easy. To the extent that the United States remains open to immigration, the availability of such labor will also help cushion the effects of an aging population.

(Data from: United Nations Population Division. "World Population Prospects: The 2008 Revision Population Database." http://esa.un.org/unpp/index.asp?panel=2: accessed August 17, 2009.)

before crises actually arise; they should never be considered as valid predictions of real future states.

Now compare across cultures, using contemporary demographic data. The United States in 2005 (Figure 2.14), although it is not pyramid shaped, still has a larger base (relatively more young people) than Europe. In part this is from the increased migration into the United States, which allows replenishment of younger cohorts. By contrast, Egypt in 2005 shows a strong pyramidal shape, heavily dominated by young people (Figure 2.15), as does the continent of Africa, even more strongly dominated by young people (Figure 2.16). This implies very different kinds of social pressure: in Europe, the problem may be finding workers; in Africa and Egypt, it will be finding jobs to keep unemployed young people from petty crime or even drifting into extremism.

The implications of such age distribution patterns for engineers are both obvious and subtle (although in most cases it is the local culture and society that the engineer should be sensitive to, not necessarily obvious characteristics such as demographic structure). Among the general observations that demographic conditions might suggest are that a young population will have different

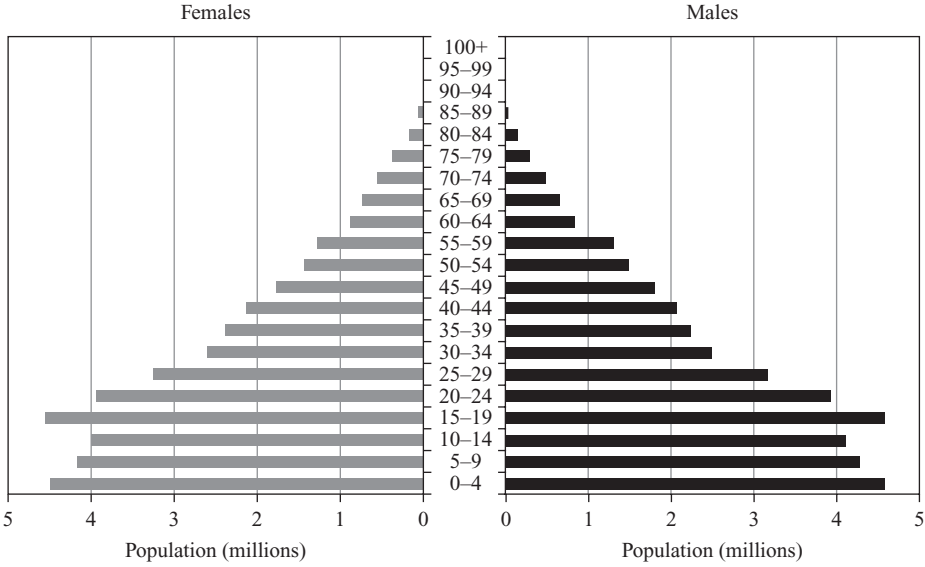


FIGURE 2.15 Population Age Distribution of Egypt in 2005. This pattern, with populations dominated by younger cohorts, is typical across much of the developing world (see Figure 2.16 regarding Africa), and differs quite obviously from the pattern in Europe and the United States.

(Data from: United Nations Population Division. "World Population Prospects: The 2008 Revision Population Database." http://esa.un.org/unpp/index.asp?panel=2: accessed August 17, 2009.)

interests and different levels of personal mobility, thereby requiring different infrastructure and building design than might be appropriate for an older population. Among the more subtle ones are that younger populations, especially where there is inadequate employment and educational opportunity, tend to be more violent and susceptible to misbehavior. This might, for example, argue that residential buildings in some regions might require more robust security in their design than in other areas.

But it is not just age distribution that is important; it is also that people over the last 100 years have been steadily living longer. Thus, even though women generally live five or six years longer than men in any given society, all individuals are living much longer than was historically the case: a thousand years ago, the average life expectancy had been estimated as about 25; by the beginning of the century in developed countries it was around 50; it is now somewhere around 70, depending on the particular country; and most experts believe it will continue to climb from there (see Figure 2.17). Initially, these extraordinary gains came from reducing child mortality; if many young children die, it reduces the average life expectancy of a population even if, once having survived childhood, an individual tends to live a relatively long life. But experts estimate that somewhere around the middle of the last century, at least

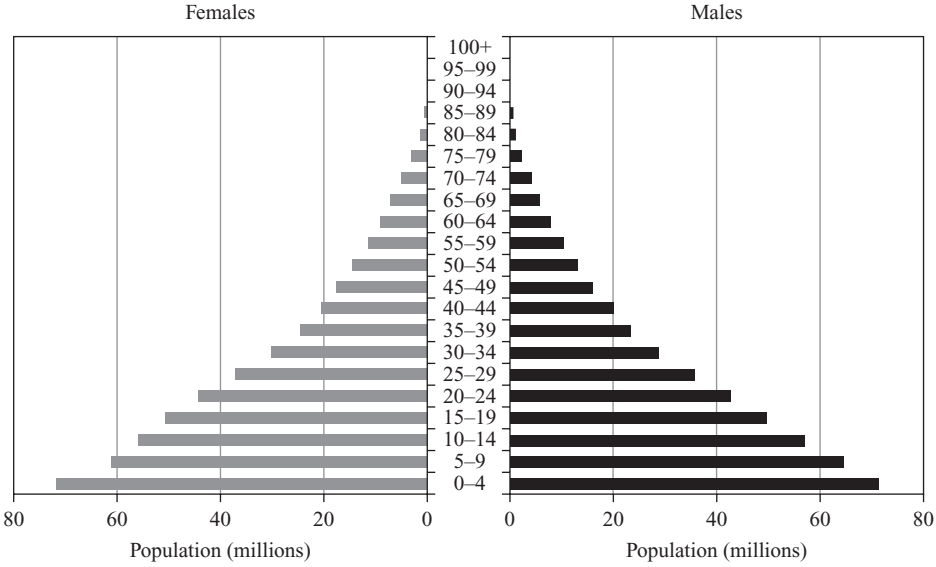


FIGURE 2.16 Population Age Distribution of Africa in 2005. This distribution is even more biased towards younger cohorts than that of Egypt (Figure 2.15). When compared with the demographic distributions of developed countries, these data may suggest that the obvious answer to aging national populations would be to allow freer immigration from developing countries, with a surplus of young workers, to developed countries, which are running out of workers. Immigration is, however, a highly charged political and cultural issue, and many countries will be unable to display the adaptability and flexibility necessary to make such a policy shift.

(Data from: United Nations Population Division. "World Population Prospects: The 2008 Revision Population Database." http://esa.un.org/unpp/index.asp?panel=2: accessed August 17, 2009.)

in developed countries, the source of continued life expectancy changed from reductions in child mortality, to higher survival rates among older individuals, so that not only were children living longer, but so were their grandparents.

Although different country conditions can reverse this growth in life expectancy temporarily (as in the case of Zimbabwe in Figure 2.17), this is a robust trend with huge implications across the board for the sustainable engineer. First, such a growth in older populations means that existing infrastructure, consumer products, and ICT services will have to be redesigned for older individuals, who may have less mobility and cognitive function than younger individuals in their prime. For example, one of the reasons that car manufacturers are working on sensor/active response systems for cars that can help avoid accidents, keep cars in their lanes, and extend perception of the automobile to its immediate environment even when driving at high speeds is that they perceive that, in the relatively near future, cars will need to be performing the cognitive functions that may be degrading in older individuals. (This area of research is called **augmented cognition** or **augcog**, and is an active engineering

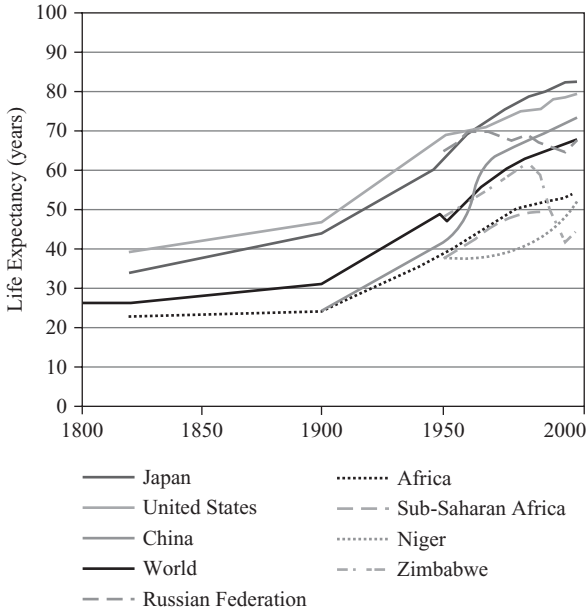


FIGURE 2.17 Life Expectancy at Birth. This graph shows the growth in life expectancy globally and for selected countries over the past two centuries. There are two major components to this trend: a reduction in infant mortality, augmented in this century by better health and longer life for seniors. Although most experts think longevity will continue to trend upwards, there are a number of interesting unknowns, such as the effects of obesity, especially on certain subcategories and regions, and perhaps most interesting, the possibility of radical life extension (to 150 years and beyond) in a relatively short time period as a result of new technology and scientific research.

(Data from: Maddison, A. *The World Economy: A Millennial Perspective*. Development Centre Studies. Paris, France: Development Centre of the Organisation for Economic Co-operation and Development, 2001; United Nations Population Division. "World Population Prospects: The 2008 Revision: Life Expectancy at Birth, both Sexes Combined [Years]." http://data.un.org/Data.aspx?q=life+expectancy&d=PopDiv&f=variableID:68: accessed August 3, 2009.)

area in military systems as well.) Cell phone designers find that enhanced functionality is popular with many people, but that there is a growing market for simple function, large numeral display, cell hones among older people for whom arthritis in their fingers and loss of visual acuity, makes iPhones and Blackberries unusable.

More subtly, an aging population combined with pension plans and retirement assumptions that are difficult to change, especially in highly unionized cultures such as in Europe, creates additional challenges. Most national pension plans are "pay as you go," which means that payments to current pensioners come from contributions by current workers. This raises obvious problems as fewer workers support more and more pensioners: The money flowing from worker to retiree simply become inadequate. The solution differs by culture.

In the United States, for example, immigration provides a continuing source of labor (as well as political conflict). In some cultures such as Japan, however, immigration is simply unacceptable as a large-scale solution for cultural reasons, so the focus is on creating robots that can take care of older individuals and do enough work in the factories to keep the economy as a whole stable. In short, the age demographics are not only driving an accelerated substitution of capital (robots and machines) for labor, but also driving the evolution of an important new technology, robotics, with huge implications for both engineering and society as a whole.

But that's not all. If societies are struggling to provide money to retirees, and if large investments in capital are required just to maintain economic stability, then it means that there will be far less money available generally for other uses. If it comes to a choice between retirees and smart grid, for example, who has the political clout? In other words, not only may these demographic trends be affecting the sorts of engineering challenges that become more common in future, but they may be constraining financing for major engineering projects across the board. Many engineers understand the importance of good technology; far fewer understand the importance of healthy financial structures to technological development and deployment. Especially in the future, as demographic changes force significant changes in design and allocation of capital, engineers will need to become far more sophisticated about the relationship between capital markets and their projects and designs.

Even these complications, however, may understate the challenge; Many scientists and engineers are working on techniques, drawing from fields as different as electrical engineering, biotechnology, and proteomics that they believe have the potential to result in dramatic extensions in longevity, with a high quality of life until very close to the end. Many of the researchers in this area believe that the first people to live to 150 or even longer have already been born in developed countries. Whether this scenario comes to pass, technological developments such as this present serious challenges to existing institutions, and social and cultural beliefs. These developments also raise obvious resource, energy, and environmental impact questions. It is important to do what we are already doing better—developing bike paths, for example. The benefits from doing so cannot be slighted. But it is also important to remember that truly transformative technologies are afoot, and it becomes an important part of sustainable engineering to design and manage them in ways that are ethical, rational, and responsible.

2.5.4 Demographics and Risk

Urbanization, shifting age structures, and other demographic changes profoundly change the risk profile of populations, in many cases in ways that have

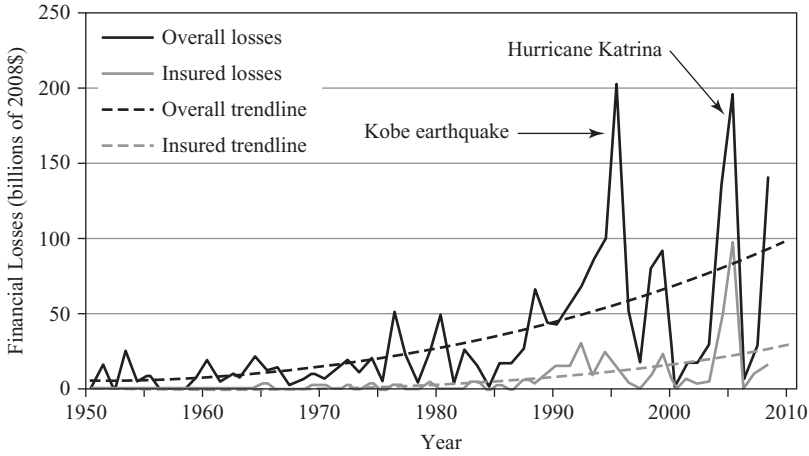


FIGURE 2.18 Approximate Global Financial Losses from Natural Disasters. The historical trend of both insured and uninsured losses from natural disasters are strongly up, showing that factors such as demographic change (e.g., increased urbanization and population growth in popular but risky coastal areas), engineering infrastructure design and construction (e.g., buildings susceptible to earthquake or extreme weather damage), and institutional weaknesses (e.g., the Hurricane Katrina response in the United States) are combining in unanticipated ways to create greater risk.

(Data from: EM-DAT: The OFDA/CRED International Disaster Database—www.emdat.be—Université Catholique de Louvain—Brussels—Belgium; Munich Re Group. "Natural Catastrophes 2008: Analyses, Assessments, Positions." http://www.munichre.com/publications/302-06022_en.pdf: accessed August 10, 2009. Reprinted with permission.)

significant implications for engineers. This change in risk is perhaps best illustrated by observing the systemic increases in damage and disruption from weather events such as hurricanes, or from natural disasters such as tsunamis, occurrences that are far more disruptive and extensive than in the past because of changing demographic patterns and the concomitant relocation of more economic activity near more risky areas such as Pacific Ocean coastlines, well-known geologically active areas (Figure 2.18).

Moreover, any such event affecting people will have differential effects depending on the relative wealth and resources of the society. As Figure 2.19 illustrates using data from the Centre for Research on the Epidemiology of Disasters at the Université Catholique de Louvain in Belgium, the countries hit by the most natural disasters are not those with the most fatalities. Because the damage done by natural disasters is not just a function of the disaster itself, but of preparedness and effectiveness of response, institutional, economic and cultural strength, in addition to good engineering, are important determinants of the human cost of disasters.

Countries Most Hit by Natural Disasters		Victims Killed and Affected, per 100,000 Inhabitants	
China, PR	35	Malawi	34,331
United States	26	Burundi	26,778
Indonesia	20	Kenya	11,935
Philippines	20	Philippines	9,097
India	17	Afghanistan	7,194
Afghanistan	13		

FIGURE 2.19 Top Five 2006 Disasters. This figure lists the five countries with the most natural disasters and the five that suffered the most human damage resulting from natural disasters in 2006 (technological disasters are not included; famines are). The only country appearing on both lists is Afghanistan, a poor country that has been in a state of war for several decades. Although the unique nature of most natural disasters and different ways of calculating how many people are affected would change this list for another index year (and mean the numbers on any such list should be taken with a grain of salt), the general pattern—that rich countries with well-designed infrastructure and resilient institutions do much better in terms of disaster management and recovery—remains valid.

(Data from: EM-DAT: The OFDA/CRED International Disaster Database—www.emdat.be—Université Catholique de Louvain—Brussels—Belgium.)

Exercises

2.1 What are the greatest differences that you see between the way you and your friends use information and communication technology (ICT), and the way your parents and their friends use ICT?

a. If you were asked to design a "simple to use and easy to operate plain old cell phone," what would you consider as your target market, you or your grandparents? Why?

b. If you were launching a new social networking site that integrated real-time digital communications of all forms—verbal, written, music, video—into unitary communication packages, what would be your target market? How would you plan to make enough money to support your service?

c. A structured document that helps a start-up business understand what its goals are, how it plans to be successful, and where it expects its money to come from is called a "business plan." Write a business plan for a proposed firm based on the social networking site mentioned in "b."

d. If you do successfully launch the site mentioned in "b," and it becomes wildly successful and widely copied, what would you expect the impacts to be on written English? Do those impacts matter and, if so, is it your responsibility to think about them?

2.2 Figure 2.18 illustrates growing losses from natural disasters. To what extent do you think this trend could be affected by different ways

of engineering relevant systems, such as buildings and local infrastructure? What do you think the barriers are to such design?

2.3 Imagine that roughly half the population between the age of 15 and 40 in affluent countries is suddenly wiped out by a very aggressive disease, for which it takes over a decade to find a cure. As a consequence, a large number of businesses will not have an educated, skilled workforce available to them and the global economy would be severely affected. What are some of the other problems that the world might face as a result? Can these challenges also be viewed as an opportunity for innovation and economic development?

a. What do you think the effect of your technologies will be on resource demand and environmental systems?

b. Is this your problem, your company's problem, society's problem, or nobody's problem? Does your answer depend on where your firm is located? Why or why not?

c. A friend knows that you are working on these technologies, and asks you whether you think what you are doing is "sustainable engineering." What is your response?

2.4 Why do you think the Industrial Revolution happened in Britain in the 1700s, and not in China in the 1300s and 1400s?

Annotated Bibliography

1. One of the more interesting books on urban systems, of particular value because of its discussion of cities throughout history, is P. Hall, 1998, *Cities in Civilization* (London: Orion Publishing). A less weighty tome, but one that covers questions of "nature" and its relationship to urban systems using the case study of New York City in a quite useful way, is M. Gandy's 2002 book *Concrete and Clay: Reworking Nature in New York City* (Cambridge: MIT Press). And for those who want to understand high modernism in urban systems from one of its most visionary proponents (you probably don't, actually), try Le Corbusier's 1929 essay, *The City of Tomorrow and Its Planning,* available in translation from the 8th French edition by F. Etchells, 1987 (reprint, Mineola, NY: Dover Publications). One of the best authors on cities as complex entities is Jane Jacobs; both her 1961 book (reprinted in 1992), *The Death and Life of Great American Cities* (New York: Random House. Reprint, New York: Vintage Books); and her 1985 *Cities and the Wealth of Nations* (New York: Vintage Books), are excellent.
2. There are a number of good, and fairly accessible, books on globalization, trade, and economic development. Among those that might interest the reader are C. Freeman and F. Louca, 2001, *As Time Goes By: From the Industrial Revolutions to the Information Revolution* (Oxford: Oxford University Press); N. Rosenberg and L. E. Birdzell, 1986, *How the West Grew Rich: The Economic Transformation of the Industrial World* (New York: Basic

Books); and P. Hugill's 1993 book on world trade, *World Trade Since 1431: Geography, Technology and Capitalism* (Baltimore: Johns Hopkins University Press). Joel Mokyr's 1990 book, *The Lever of Riches* (Oxford: Oxford University Press), focuses on the role of technological creativity and its role in economic progress, and thus might be of particular interest to students of sustainable engineering.

3. J. Diamond, 1997, *Guns, Germs and Steel* (New York: W. W. Norton & Co.) and D. S. Landes, 1998, *The Wealth and Poverty of Nations* (New York: W. W. Norton & Co.). If you're going to work with technology systems and sustainability, it's important to understand why we've ended up where we are now. Why, for example, aren't most people speaking Chinese rather than English? These books provide interesting perspectives on this most difficult problem. Some people see them as contradictory—Diamond viewed as "leftist" and Landes as a "rightist"—but I think they are best read as complementary. And both address an absolutely critical environmental question: Why does the modern world look the way it does, with a globalized, Eurocentric Enlightenment culture? After all, if you don't understand at least a little about why the world looks the way it does and what the critical dynamics are, how do you think you're going to change it?
4. For those who want one relatively intelligent and reasonable weekly source of news and commentary, it is hard to go wrong with *The Economist.* The writing is excellent and generally aimed at a more educated audience than competing news weeklies; the advanced economics that does pop up in some articles can usually be safely ignored. Moreover, if you read *The Economist* on a regular basis, you can begin to develop a feel for many of the trends discussed in this chapter, and others that haven't been covered, for yourself. But like many English imports, it can be an acquired taste, so, given that it's not cheap, it's probably worth buying an issue or two at the news stand and seeing if it appeals.

CHAPTER

3 Complexity

. . . in our current world situation, opening oneself or one's group to a larger "database" reveals the terrifying prospect that the world is now so complex that no one really understands its dynamics and that even rational efforts tend to be washed out or misdirected by processes not understood and consequences not anticipated. Of course, as suggested earlier, those intent on pursuing their interests seldom can risk sociocultural ostracism by acknowledging this to others, and usually not even to themselves.[1]

—D. N. MICHAEL

[1]D. N. Michael, "Barriers and Bridges to Learning in a Turbulent Human Ecology," in eds. L. H. Gunderson, C. S. Holling, and S. S. Light, *Barriers and Bridges* (New York: Columbia University Press, 1995), pp. 461–488, 473–474. This is a notable observation, in part because Michael has spent much of his time trying to understand and manage complex ecological systems, which necessarily entangles human, built, and natural elements—precisely as the anthropogenic world does.

3.1 COMPLEXITY

One theme of the anthropogenic Earth that is important enough to sustainable engineering to deserve its own chapter is complexity: complexity of technology systems, of both designed and cultural information systems, of the accelerating integration of human, natural, and built systems that characterizes our time. The usual tendency, especially in public dialog, is to reduce complexity to slogans and simplifying symbols: Carbon footprints, Islamic terrorism, and environmental scorecards are examples. But although the sustainable engineer may often work with stakeholders, institutions, methods, and tools that simplify complex systems, he or she needs to develop a more sophisticated understanding of, and embrace, complexity. It is not that simplicity is wrong. Indeed, it can be very useful, even necessary, especially when communicating with the public. The key is in knowing what can safely be simplified, given the particular problem the engineer is challenged to solve.

The challenges of complexity can be seen in many ways. For example, it is far more difficult than it used to be to design, build, and operate mining or oil production facilities in developing countries. This is not just because the internal politics and institutional relationships involved are more complicated—think of the conflict between communities, brigands, and the Nigerian state in the oil-producing Nigerian Delta—but because in many instances human rights and environmental NGOs will be actively intervening, either locally or by initiating litigation in many countries regardless of where the project may be located. This is institutional and ideological complexity. Alternatively, designing a new search engine is not just about good software production anymore; especially as Internet services become increasingly challenged by national laws and the desire of some countries to maintain strict censorship regimes, it becomes a delicate balance between individual rights, and state policy and security, as well. Nor is this complexity always amenable to rational optimization. For example, the Tata Nano, an inexpensive automobile that promises to dramatically expand automotive use in countries such as India, has been widely praised by local consumers, for whom it offers a much safer and efficient means of transportation—and just as widely denounced by environmentalists from developed countries because it contributes to fossil fuel use and global climate change. Who's right? There's no objective answer that will satisfy all stakeholders. In such cases, then, a good engineer is one who can develop such projects in a technically sound way, while also minimizing the conflict with the increasing number of stakeholders involved in the process and respecting relevant cultural constraints and ethical precepts. But to do so requires understanding and accepting complexity, not denying its existence.

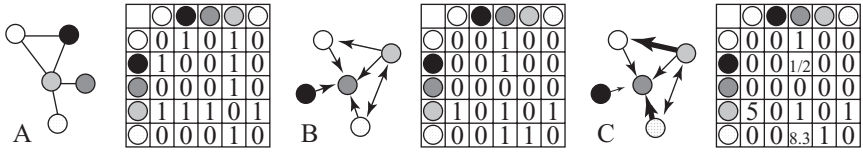
3.2 SIMPLE VERSUS COMPLEX SYSTEMS

Let's begin by providing a few simple definitions. **Systems** are simply groups of interacting, interdependent parts linked together by exchanges of energy, matter, and/or information. Although this sounds straightforward, it hides a very important principle that technologists should always be aware of: Almost all systems, and all complex systems, depend heavily on what boundaries are chosen to separate what is considered to be part of the system and what is considered to be outside the system. Drawing such boundaries is a matter of judgment and art, and it is always a somewhat arbitrary action. This means that results of systems analysis can be biased, or even predetermined, by selecting different systems boundaries.

Consider, for example, the simple question of what the environmental impacts of jet airplane technology are (or in systems language, what is inside the system of environmental impacts of jet airplanes for the purposes of a lifecycle analysis, and what is not). Most people immediately recognize that fossil fuel use, and perhaps the environmental impacts of materials used to build jets, should be in the system. There is less agreement on whether the infrastructure necessary to support the technology—airports, parking lots, energy used in air traffic management operations, and so forth—are included. Finally, most people would not include the well-known effect of airplanes in spreading disease vectors, or enabling ecotourism, which degrades previously unreachable ecosystems, or adding significantly to the quality of life of senior citizens by providing opportunities for inexpensive travel. There is no "right" definition of the boundary, and thus the relevant system, for such an analysis. Rather, the appropriate boundary depends on why the analysis is being done and what the results will be used for.

Note some of the implications of this unavoidable ambiguity of systems boundaries. First, it puts a substantial burden on the person doing the analysis—which may well be you—to be responsible and ethical in defining the relevant system reasonably. Second, it means that if somebody wants to predetermine outcomes, and do so without being detected, that person doesn't need to be explicit or falsify data; the person can simply choose his or her boundaries in such a way as to produce the desired outcome. You should not do this, and you should be very skeptical of any analysis that does. Finally, it means that you need to be careful with the application, especially under novel circumstances, of any systems analysis method or tool, such as lifecycle assessment tools available on the Internet. This is because all such tools have to begin with some framing of what will be considered, and what won't—some systems boundary, in other words. There is no reason to assume that such a generic definition will necessarily be appropriate under new conditions.

Systems vary in fundamental ways. For example, Figure 3.1 shows three schematic systems, structured as networks (an increasingly common way to



Simple Graph: Symmetric, Binary.

Directed Graph: Non-symmetric, Binary.

Directed and Weighted Graph: Any Matrix.

FIGURE 3.1 Different Types of Networks. Network A is characterized by undifferentiated interconnection between particular nodes. Network B is similar, but now the linkages flow in specified directions, as, for example, a simple supply chain feeding into a factory might. Network C illustrates a further differentiation, in that some of the links are now identified as more important (are weighted more heavily) than others. Again using the supply chain example, some materials are critical to production and would thus be heavily weighted linkages, while others—detergent to clean machines, for example—are part of the supply network of a factory, but much less important, and thus much less lightly weighted. Note that weighting in such cases can reflect many different parameters (volume of material, importance to production, or something else), and thus need to be specified as part of the network's definition.

Graphic provided courtesy of César A. Hidalgo.

think about networks and, as the figure shows, a good way to move between matrix and visual presentations of systems structure).

Differences among systems may have profound real-world implications. For example, some systems may be heavily interconnected, with almost all nodes tied to many other nodes, while others may lack connections almost entirely. In the first case, evolution may be very hard, because a change in any single node is quickly damped out by all the other nodes it is connected to. Thus, for example, it is relatively easy to design a car that runs on hydrogen, but one cannot evolve to a national fleet of hydrogen-powered cars until a suitable energy infrastructure that can deliver hydrogen as easily as we now deliver gasoline can be developed. Similarly, one cannot introduce plug in vehicles until the electric grid is evolved to manage the increased load. In each case, the implementation of the immediate technology—hydrogen or electric vehicle—is dependent on coevolution of a complex set of other technology systems to which it is tightly coupled. It is an important skill for sustainable engineers to be able to understand, and work with, such linkages.

Conversely, however, a complex system with a few linkages also has trouble evolving. Even if a node changes, the lack of linkages to other nodes means the system as a whole doesn't evolve very effectively, so one tends to get confused localized behavior rather than system evolution. History is replete with interesting inventions such as Charles Babbage's famous Analytical Engine designed in the 1840s, a mechanical calculator that could not be built because the materials and manufacturing skills necessary to do so did not yet exist. Obviously, the best system for adapting gracefully to

changing conditions is one that has enough interconnectivity, but not too much, so that the system can hold together as it evolves, but not be too constrained by its own internal structure.

For the sustainable engineer, there exist tools to help achieve such a state. The electronics industry, for example, has jointly developed a set of technology road maps that provide some framework for evolving technologies and systems, without overly constraining the specific products and processes that various innovators will develop. (Road maps also provide a good example in other ways of the challenges of sustainable engineering in complex technological systems: An important element of any such road-mapping activity is making sure that one complies with applicable antitrust and other legal regimes.) In other cases, constraints will be apparent and integrated into design almost unconsciously. Engineers don't design toasters that run on random voltages, or cars that are one and a half lanes wide, or trains that run on arbitrary gauges (the size of the space between rails). Especially in regard to technology, complexity in these systems may also be **wicked**, involving social, cultural, and institutional constraints (discussed in more detail below). Getting new technologies smoothly to market doesn't just depend on how they couple to other technology systems, but whether the innovator can find financing, easily establish a new company, protect intellectual property claims, take risks of failure, depend on the legal system to establish defensible and nonarbitrary contract rights, is respected by society, and the like.

How a system is structured also affects how resilient it may be to changes in its environment. For example, so-called scale-free networks, in which one finds a few superconnected hubs, behave quite differently than ones where connectivity is just random (Figure 3.2). Scale-free networks are quite resilient to random failure, precisely because if a link or hub goes down, chances are that the network structure can simply route around it. On the other hand, such networks are quite sensitive to targeted attack; take out a superhub or two, and you bring down a scale-free network. Is this relevant to the real world? Given that many critical infrastructure systems are scale-free networks—absolutely! The Internet, for example, is very resilient to random failure, because it is a scale-free network. For example, during the September 11 attack on the World Trade Center, Internet technology generally continued to work quite well, because although it was an intentional attack, it was not an intentional attack on the Internet, which was thus quite resilient. On the other hand, this also means that, if one wants to take down the Internet, a targeted attack, if it can get through to the superhubs, may be quite successful. (This is the basis of so-called "denial of service" attacks that use pirated machines—bots—to launch many messages to one set of servers, thereby overwhelming them.) These systems characteristics are obviously basic considerations if you are, for example, designing software or Internet security, or working with military information systems that must be kept operational. Increasingly, virtually all infrastructure

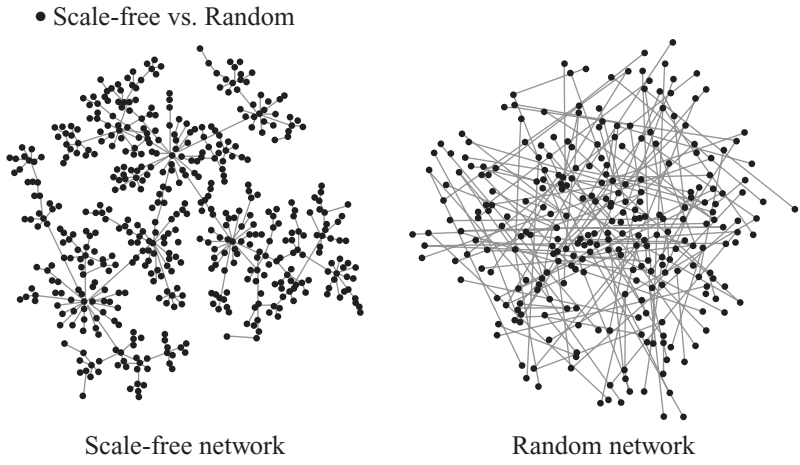


FIGURE 3.2 Scale-Free and Random Networks. Network graphics are used to illustrate two kinds of complex systems: a scale-free network and a random network. The operating implications of each are very different, which is why an engineer needs to know not just whether the system he or she is working with is complex, but what kind of underlying structure, and thus behavior, is involved.

Graphic provided courtesy of Ming Xu.

systems, military or civilian, integrate ICT capability in such ways that attack can occur either physically, or in cyberspace. For instance, an assault in 2007 on Estonian Internet functionality by Russian elements as part of a political dispute didn't require physical destruction of key superhubs (which might well have been considered an act of war), but simply focused attacks on core functional components in cyberspace, which were just as effective, but important in an international context, allowed the Russian government to deny responsibility (whether they were responsible or not remains technically unclear). In that case, both private industry, and military and governmental, Internet systems were targeted, and both went down temporarily under the stress of the attack.

As this discussion suggests, it is useful to differentiate **simple systems** from **complex systems**, because their implications, and even how well they can be understood, depend heavily on this difference. (Table 3.1 summarizes some of these differences.) In contrast to simple systems, complex systems are characterized by strong and usually nonlinear interactions between the parts. Moreover, complex systems contain feedback loops that make it difficult to distinguish cause from effect. They also display significant time and space lags, discontinuities, thresholds, and limits, and tend to operate far from equilibrium in a state of constant adaptation to changing conditions. (As complexity scientists say, they operate at the edge of deterministic chaos.)

What does all this somewhat general language mean? Consider two examples. The first is a salt marsh that, because of poor storm drainage design

TABLE 3.1 Simple (S) Versus Complex (C) Systems[1]

Function	Type	Function as Displayed by System	Policy Implication	Example
Information	S	System is centralized and "knowable"	Centralized command-and-control feasible	Environmental end-of-pipe regulations
	C	Information diffused throughout the system; some embedded in system structure; system too complex to be "known"	System not directly manageable; real-time adjustments to unpredictable behavior, not centralized control	Climate change management
Causality	S	Linear; cause and effect relationships easy to see	Centralized command-and-control to end point (effect) feasible	Removing lead from gasoline leads to less lead exposure
	C	Causes and effects cannot be linked in most cases	Cannot be sure of actual impact on system of any policy initiative	Designing the Everglades; running modern large city
Response to Forcing	S	Predictable and relatively linear	Rigid regulatory structures both okay (because outcome predictable) and politically preferable (because can't be gamed)	Removing lead from gasoline
	C	Highly nonlinear and may be discontinuous; not predictable a priori	Policy once in place must be adaptive and flexible, so can be changed in real time as system's response dictates	Designing Everglades; climate change management
Control Mechanisms	S	Rational centralized control is possible and effective	Single, fully responsible entity with authoritarian power can control system (e.g., U.S. EPA)	Legal controls on interest payments
	C	Rational direct control fails because of lack of knowledge and system unpredictability	Responsible management entity adjusts forcing functions (e.g., taxes or fees) and monitors results; system has diffuse internal control mechanisms	Regulating modern economy versus Soviet command economy
End point	S	Can be defined and not path dependent	Command-and-control management to end point feasible	Constructing a bridge or building

TABLE 3.1 (Continued)

Function	Type	Function as Displayed by System	Policy Implication	Example
	C	System has no defined end point or equilibrium point, and must be understood in process or performance terms	System can be managed to have desired behaviors emerge, cannot be managed towards a defined end point	Adaptive management of forests or fisheries; engineering the Baltic Sea
Metrics	S	May not be necessary because performance measured by attaining end point, not monitoring system state	End points are metrics	Successful completion of an engineering design or project
	C	Defined in process or performance terms, and change over time because system itself constantly evolving	Manage process adaptively over time to improve performance against metrics; choice of appropriate set of metrics essential, but subject to improvement over time as well	Engineering the Everglades; designing climate change adaptation systems
Existence	S	Produce end point	Can minimize specific insult (e.g., lead dispersion from use in gasoline); cannot lead to or support achievability of sustainability	Chemical and material bans
	C	Maintain system stability and integrity over time; build systems functions such as resiliency	Policies with complex systems must be adaptive and focus on behavior of system as a whole. Managing to end points does not work; rather, goal must be on maintaining system performance given unpredictability of system	Engineering the Everglades; designing and managing infrastructure in a large city

[1]This table illustrates some of the major differences between simple and complex systems, and the policy and design implications that flow from those differences. The primary difference is that simple systems are, under proper conditions, understandable, predictable, and thus manageable directly; complex systems, on the other hand, cannot be intuitively understood, are inherently unpredictable, and cannot be managed in any direct way. Put another way, simple systems to some extent can be engineered directly; complex systems may require nontraditional engineering approaches, such as continuing dialog with an evolving system rather than simply designing an artifact and walking away.

in a nearby town, is exposed to increasing levels of pollution. The salt marsh appears not to change much for a while, but as the health and numbers of a key species of salt marsh grass drops below a critical point, it quickly degrades into a denuded mud flat. This is a nonlinear response—a small, additional change results in a dramatic shift in system state. (This nonlinearity is familiar to us in proverbs such as "the straw that broke the camel's back.") It also illustrates feedback loops. The damage to salt marsh grass does not stop with the grass, but feeds back into other species, reducing them and inducing major system collapse. (In fact, ecologists call species that have such critical roles in ecosystems "keystone species.") There are discontinuities and lags as well: The jump from healthy-looking salt marsh to mud flat is not a smooth change over time, but a rapid shift to another, very different state. Finally, any complex system like a salt marsh is never in a "final state," but is always adapting to new conditions. Although we may not like it as much, one way to understand the ecosystem of a denuded mud flat is as an appropriate adaptation to increased pollution loading.

The second example involves stock trading on Wall Street, where the ability to make trades rapidly to capitalize on temporary small differences in prices (arbitrage) can make lots of money. Starting in the 1980s, as computers became more powerful, more and more stock traders began to mechanize their trading so that they could capture profits from arbitrage before others did. Their trading programs frequently had "floors" built into them, so that, if the market began falling, the computer began selling rapidly to avoid taking too large a loss. The first few adaptors of this technology began making large profits, so of course the technology spread rapidly. The stock market always fluctuates, however, and in October 1987, for a number of reasons, none of them particularly critical in global financial terms, it fell relatively rapidly to a point where some of the trading programs began automated selling. Selling stock drives the market down, so this programmed selling drove the market down further and faster, which led other programs to initiate selling. Eventually, these events cascaded into a major stock market collapse. There was no external financial crisis that justified this collapse; it was largely a result of internal system structure and dynamics.

This is a nonlinear response, because in the critical region of the system (that is, in the price range where the computers began programmed selling) a small change in price induced a large change in system behavior (everyone started to try to sell, and as traders say, "The bottom fell out of the market"). The feedback loops are apparent: Changes in one computer's behavior rapidly rippled across other computers. The threshold effect is also apparent: In price regions well above the critical price region, the computers functioned as intended, but when the price hit the programmed selling point, the drop in

the market was immediate and discontinuous. Finally, as the subsequent panic of 2008 illustrates as well, there are no equilibria in the stock market, only complex adjustments as stock prices, and the market itself, changes in response to internal and external dynamics, including such nonequilibria factors as consumer psychology and the state of international conflict.

These examples illustrate another important point. Many systems, whether biological, economic, technological, or cultural, may behave as if they are simple under some circumstances and may behave as very complex systems under others. Thus, systems engineers might say that a system is in its **linear regime** when it changes predictably and smoothly in response to small changes in its environment, but in a **nonlinear regime** when it changes unpredictably and dramatically in response to a small change. In addition, unlike complex systems, simple systems are relatively easy to manipulate and frequently characterized by weak linkages among components. This means that they are less likely to be irreversible and illustrate "path dependence." For example, if one is making a fruit salad (a simple system), one can put in the cut apples after or before the melon balls, and it won't affect the outcome; moreover, the apple slices can be removed quite easily if you don't happen to like them. But if one is making a chocolate cake, it is quite impossible to remove the egg or the flour after it is baked.

Part of the dynamics of complex systems is that they evolve, and that evolution is path dependent. What this means, simply, is that complex systems respond to changes in both internal and external environments by changing states in ways that enable them to continue functioning as much as possible, but that they can only adapt as much as their current structure allows. This requires some mechanism in the system by which information is captured, stored, and transmitted; it requires that the system is able to generate new states, some of which will be more adaptive to new circumstances to others; and it requires selection among alternatives based on performance given internal states and external boundary conditions. Once a cake is made, for example, you can't go back to the ingredient list; you have to go forward, and frost and eat the cake. To take a more complicated example, a beetle cannot suddenly develop wheels; its biological and genetic pasts have created a path where legs can be evolved, but wheels can't. Similarly, it is impossible to develop a technology that is completely different from all that went before, because our educational system, our training, and the technological infrastructure preclude it.

But, you may argue, we develop new and unpredictable technologies all the time. Of course. The trick with complex systems is that, although specific aspects of their evolutionary paths are unpredictable, the general space within which that path lies is established by the past history of the system (as illustrated

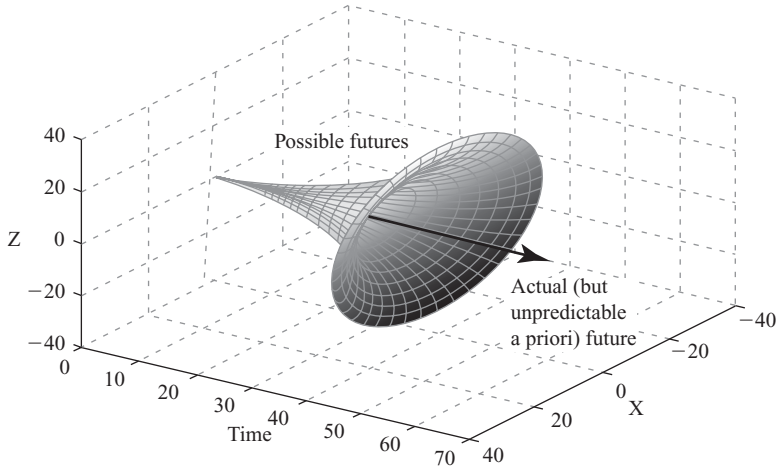


FIGURE 3.3 Path Dependency. This figure illustrates the path dependency, but unpredictability, inherent in the evolution of complex systems over time. In most systems, of course, the space involved is not a three-dimensional one, as illustrated here, but a many-dimensional one (even defining the axes of such spaces can be a challenge).

in Figure 3.3). To return to our beetle: Whether the beetle evolves black, brown, or red legs may be unpredictable, but that it will not develop wheels given its current genetic and biological structure is apparent. Or to take another example, it is impossible for me to predict with any certainty whether it will be rainy or sunny in New Jersey on July 4, 2020, but I can predict with a fair amount of certainty that it will not be snowing. In other words, path dependency combines with unpredictable specifics.

Why does complexity matter? Simply put, it's because most engineering involves complex systems. But from a sustainable engineering perspective, it's also important to realize that the kind of complexity involved in many engineering activities has changed not just in degree, but in kind. Designing the Everglades is not just like designing a toaster, only bigger. It requires a very different idea of what engineering, design, and management of the system in question is all about. For example, it used to be that if you were the project manager building a road into a remote Amazonian location for an oil well, you merely had to worry about building a good road (well graded, good drainage, bridges designed to carry heavy loads, minimum transport of fill, and the like). Now, however, you need to worry about how your road may affect indigenous native peoples, create corridors for hunters and slash-and-burn agriculturists, degrade local environments, and stand up to court cases brought in the United States by activists determined to stop your project. Good design, construction, and operation are still the goal; it is just a lot more complicated to do in practice.

3.3 DIFFERENT FORMS OF COMPLEXITY

In considering complexity, it is useful to break it down into four related but conceptually separate domains. The most obvious and easy that engineering almost always engages directly, is static complexity. Static complexity arises from the number of things and their connections that are involved in a particular system. For example, an automobile or airplane simply sitting on the ground is a highly complex object in part because of all the parts, subassemblies, and physical and information linkages from which it is composed. A computer chip is highly complex in part simply because of all the operational devices of which it consists. Moreover, as we demand increasingly challenging functionality from our products, their static complexity increases, as Figure 3.4 illustrates.

The second domain is dynamic complexity, which looks at systems as they move through time and respond to changes in external and internal states. Thus, for example, an airplane consists of many interconnected parts (static complexity), and these parts must interact in predictable (and safe!) ways as the airplane actually functions (dynamic complexity). **Dynamic complexity**, however, does not necessarily require **static complexity**. Business school students learn this when they play the "beer game," in which they try to balance demand and supply in a simple model that has a producer, a wholesaler, a retailer, and

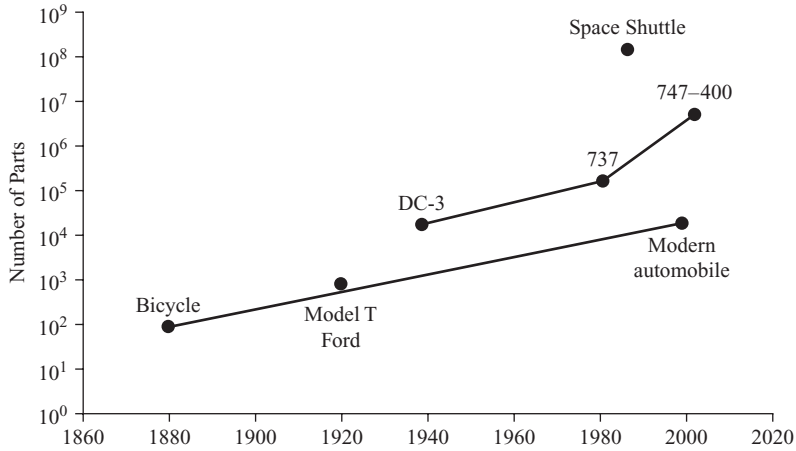


FIGURE 3.4 Static Complexity Over Time. This graph illustrates how the static complexity—the number of parts per product—has tended to increase over time. It is a testament to modern engineering that, despite the dramatic increase in static complexity, such products as modern jet aircraft can function for thousands of hours without a failure endangering the consumers relying on it.

(Adapted from a sketch provided by Paul Sheng, University of California, Berkeley. Printed with permission.)

critically, lag times among those three. Although this sounds trivially simple, in almost all cases students will crash the system, because its dynamics, almost entirely a product of the introduction of lag times between nodes, are far more complex than the static model would imply. More broadly, agent-based modeling has shown that quite complex behaviors—fish schooling, for example—arise from a relatively few agents (e.g., fish) and a few simple rules. Engineers are now applying this lesson in many interesting ways, such as developing insect robots that are capable of quite complex behavior based on a few simple rules.

Engineering and science students are generally familiar with these two forms of complexity, which can be quite complicated but are approachable using quantitative tools and methods. The third type of complexity, however, is one with which they are generally not familiar: **"wicked" complexity**. *Wicked* here is a term of art; it means systems that include important human components. Such problems have a number of important characteristics. For example, they cannot be definitively described, and any description is always partial. It is also not possible to reach any objective definitions of the optimal system state, because concepts that people define quite differently, such as *equity* or *fairness*, are important determinants of how a particular state is evaluated. Policies and designs for wicked problems cannot be meaningfully correct or false. Finally, there are no optimal solutions to wicked problems, only workable ones—and workable means not just technically feasible, but politically, socially, and culturally acceptable to the relevant stakeholders.

What are some wicked problems? Suppose, for example, you are one of the engineers chosen to help manage the Florida Everglades. You immediately find that there are a number of completely different worldviews at play: developers may feel that water for new communities is most important; environmentalists will want to stop development to preserve water for the Everglades ecosystem; farmers will want to continue growing sugar cane, which in turn exists in Florida only because of trade policy imposing import barriers on cheaper Caribbean sugar; Native American tribes will want to protect existing ventures such as casinos; and money to fix problems is always limited and sometimes directed by specific interests. None of these positions, which are based on different sets of values, are demonstrably "right" or "wrong"; as the engineer who must try to balance them in the real world, you cannot overly favor any particular perspective (optimize on environmentalism, for example) because the result would be politically unstable. Moreover, the balances and appropriate design measures may well change over time: a housing bust in Florida may reduce the importance of development, or a sugar farmer may offer to sell his land, creating an opportunity to improve environmental performance of the system as a whole with less political friction.

Similarly, suppose you are the engineer in charge of running a major mining operation in an unstable developing country. You will find that some sustainability advocates hold that any mining is unsustainable, and therefore demand that you shut down all operations immediately. Your shareholders and managers want to maintain a flow of profits. Local workers may want to retain their employment. Western environmental organizations may be concerned with the environmental impacts of operations. The national government may take all the returns from mining to line their own pockets, leading local communities to attack you to obtain what they regard as a fair share (something like this happened in the Nigerian Delta regarding oil operations). European and American NGOs may demand limits on design and operation of the mine that make it unprofitable, or suggest worker and community programs that the local government finds unacceptable, and so forth. There is no "right" answer that meets all of these conflicting demands and worldviews; there are only better and worse ways of trying to muddle through in a rational way. For example, environmental efficiency in operation is, all things equal, always a good idea, but if you work for the mining company, or consume products that use metal, shutting down the mine in the name of sustainability is not going to be feasible.

The point is not that wicked systems are inherently irrational. They are not. They had better not be, because they are increasingly the type of systems within which engineers have to work. They are, however, complex in ways that most technologists and engineers aren't trained for. In fact, one reason for this book is to provide some exposure to such systems, so that when you have to grapple with them, you know what they are.

A final level of complexity arises from the scale of human activity in the Anthropocene. As the last chapter discussed in more detail, this complexity involves wicked systems at regional and global scales; the interplay of technology, ideology, partial information, differing worldviews, and dramatically different distributions of costs and benefits of technologies, economic initiatives, and policy decisions within and between cultures makes engineering at this level fraught with difficulty. If you design the Tata Nano automobile, are you socially concerned and aware (reducing unequal access to automotive technology and building safer transportation alternatives), or environmentally evil, as many Western NGOs charge? If you build a road into the Amazon to facilitate agriculture, are you doing something worthwhile, because it enables production of more biofuel and food, or bad, because it enables exploitation and degradation of the rain forest? If you design the perfect "plug-in" electric automobile, are you really doing sustainable engineering (see the discussion in Box 3.1)? Especially because a globalized economy tends to scale up technologies very quickly if they are successful, the earth system level of complexity, although still inadequately recognized, should always be in the back of the sustainable engineer's mind.

BOX 3.1 Plug-in Automobiles and Biofuel: Why Not to Fall in Love with a Technology[2]

People have an unfortunate tendency to fall in love with certain technologies. Several years ago, it was hard to find anyone who had a bad word for biofuels. They were "renewable," "natural," and "sustainable," and were widely embraced by activist and government regulator alike as the means by which a nasty fossil fuel based economy could be transitioned to a more moral and righteous state. In short, everyone fell in love with them. But love is fickle, and nothing is harder on fantasy than reality, as every engineer soon learns. So countries such as the United States introduced biofuels rapidly and on large scale, and suddenly faced emergent behaviors that everyone had forgotten about: increases in food costs, biodiversity impacts of increased demand for agricultural land, increased demand for water—and we're still assiduously ignoring accelerating perturbations of other critical systems such as the nitrogen and phosphorus cycles. Moreover, because the relevant markets are global, so too the impacts. European and American policy decisions rapidly ramping up demand for biofuel do not just affect those countries' citizens, but the food prices and well-being of people around the world, cruelly, especially the very poor. These are complex systems, and pretending they are simple doesn't make them so.

The critical lessons, although often overlooked, are not complicated. Most important, don't fall in love with anything—not an ideology, not a buzzword like *renewables* (we're still in love with that one), not a technology. Anticipate that any system worth considering, such as biofuels, exhibits nonlinear behavior, especially if scaled up rapidly without allowing for adjustments throughout all the other systems coupled to it. At some scale, such technologies are generally good ideas; at large scale, the nonlinearities are quite likely to cause ugly complications, especially if demand is not eased in, but spiked (corn-based ethanol being the current example). Moreover, always remember that technologies over time become technically, economically, and politically locked in; trying to change America's corn ethanol policy is far more difficult now that Midwest farmers, construction workers, and capitalists see the potential benefits. These farmers have changed their farming patterns, and the constructions workers and capitalists have begun to build corn ethanol infrastructure—all three stakeholders have a strong and vested interest in continuation of the program.

And yet, we're doing it again. The hysteria over climate change, combined with the success of hybrid automotive technologies, is generating huge pressure for "plug-ins," electric automobiles that can be recharged from the grid, requiring no petroleum at all—locally, at least. Such vehicles are seen by the public, and represented by their proponents, as "emissions free," "carbon neutral," and "important moral statements," and are otherwise worthy of undying love. Environmentalists and progressive regulators are increasingly demanding that we stop half-measures that continue to destroy the climate system, such as hybrids, and go directly to plug-ins. Such a

[2]This box, as occasionally elsewhere in the text, is modified from columns done over the period 2007–2009 for the online publication *GreenBiz* by the author.

policy may work. Most likely, especially given the lack of systems understanding and analysis, it will be corn ethanol déjà vu.

There's no question that plug-in technology is cool. Oh, there's specific technology issues, such as finding batteries appropriate to automotive power demands and use patterns; figuring out viable combinations of on-board power and rapid recharge technologies; and creating a viable charging infrastructure (e.g., the smart grid). But the idea of using clean electricity to phase out one of the most obvious and intractable uses of fossil fuel, without having to do the difficult job of weaning people off cars completely, seems too appealing to pass up.

There's much bigger problems with this love affair, however, and they're not being discussed much. What happens if the activists succeed, and we suddenly lurch towards plug-ins at large scale? Remember that there's nothing inherently clean about electric power, not today. So what kind of electric power production and distribution infrastructure does a large demand spike imply? Well, coal plants are the most likely and cheapest addition to base load capacity under such a scenario of rapid demand development. At the least, plans to take older, more polluting plants out of production would have to be shelved. It's highly unlikely that wind or solar can substitute in the short-term. Most people don't realize the difference between existing technologies, such as coal, nuclear, and natural gas, which provide reliable power generation, and many forms of renewables, which are variable depending on the source. (For that reason, Britain's National Grid estimates that it takes 25 gigawatts of wind power production capability to replace 5 gigawatts of fossil fuel fired production.) In short, renewables are a long way from being able to substitute for traditional fossil fuel production, especially in the short run. So, with cars evolving over a couple of years cycle time, and the energy production infrastructure evolving over decades, a sudden spike is much more likely to lead to construction of more fossil fuel plants. This is especially true as countries such as China and India, already struggling to generate sufficient electricity for their development, face yet more demand for electricity from plug-ins.

It's not that plug-in technology may not be useful at some scale. It's that, by falling in love with the technology, we once again blind ourselves to the questions we need to ask to manage a complex engineering transition. And we're overlooking the same old things: What are the coupled technologies, and what are their cycle times (in this case, electric energy production and distribution infrastructures)? At what scale can we introduce the new technology, so we can learn about it, without pushing coupled systems into nonlinear behavior? How fast can the transition be, and what resources will it require? For example, private car firms can certainly build plug-ins; the "smart grid" and increased energy production capacity they imply, however, will require a lot of financial capital, and it's not clear where that will come from, given the debt load that many countries have assumed to fight the panic of 2008. And by what metrics can we steer this technological innovation to assure that some of the desired social benefits—producing fewer carbon emissions per unit quality of life, for example—are met, rather than the exact opposite—a new technology, introduced at large scale without thought, driving increased installation of fossil fuel power generation capability?

If you want to be a sustainable engineer, by all means fall in love. Just don't do it unthinkingly with a technology.

3.4 UNDERSTANDING COMPLEXITY

This discussion may have made it sound as if we understood complexity. We really don't, as a couple of cases readily suggest. Consider, for example, the stock market collapse brought on by the computerized trading programs mentioned earlier. This involved something that people care a lot about—money—and a mechanism of system failure—computer sell programs interacting in a dramatic feedback loop—which is hardly rocket science. Yet the market collapse occurred despite what, in retrospect, was a blindingly simple system dynamic. More recently, the financial panic of 2008 involved conceptually similar issues, as the complexity of financial instruments outran the ability of not just regulators, but even of the people putting them together, to understand them; the result was that, when financial markets began contracting, no one really knew where the risk was, or what many of the financial derivatives were actually worth. In many ways, the 2008 incident illustrates the dangerous mix we have today: We're working with far more complex systems than ever before, but our understanding of them hasn't kept pace with their evolution. Don't forget that the people who almost brought you a second Great Depression were called—wait for it—"financial engineers." Indeed, many of the people on Wall Street, and in global financial institutions, were originally trained as engineers.

Another example of mishandled complexity played out in the great confrontation between capitalism and communism. Communism in the Soviet Union and China collapsed not from external conquest or even primarily from the military spending race that the United States allegedly launched. Rather, the centralized economic model adopted by large marxist societies essentially became incapable of managing the complexity inherent in a modern industrial economy. It was not that central planning on the Soviet model was completely inadequate; the Soviet Union developed very rapidly in certain areas, such as military and space hardware, although towards the end the military production sector was having trouble keeping up with the United States. But it failed across the broader, more complex economy. At one point, the Soviet Union ran the only significant economy that was capable of taking in large amounts of raw materials and producing products from them that were, in aggregate, worth less than the original raw materials—a production system that subtracted, not added, value. This was because the Soviet Union made a fundamental conceptual error: It treated a complex system, the modern economy, as if it were a simple one, easy to understand and to control centrally (see Table 3.1, above). Interestingly, partially because China never had the actual control over its economy that the Soviet Union did, it escaped the systemic collapse that the Soviet Union suffered. For the sustainable engineer, this suggests an important lesson: Managing complex systems requires relying on indirect and dispersed information and control mechanisms, as in a market economy, rather than on

the centralized and explicitly defined and controlled systems that some engineers are more comfortable with.

Along these lines, it is important to realize that, with these sorts of systems, any understanding we achieve is usually partial and based on the purposes for which we are working with the system. Consider a simple example involving cities. If I am interested in the rates of crime in the city of Phoenix, I am also implicitly defining the urban system by its political boundaries, because most crimes are local and statistics are kept on a jurisdictional basis. If, on the other hand, I am interested in water supply and Phoenix, thinking about "Phoenix" as only the geographic city is not sufficient. By my question, in fact, I am implicitly defining the system as including the Colorado River basin, not to mention American water law, patterns of tourism that make golf courses popular in Scottsdale, and xeriscaping initiatives. Yet in both cases the relevant marker is "Phoenix." Neither way of thinking about Phoenix, or indeed the many others that a minute's thought will suggest, is right for all purposes. The system itself always remains more complex than the sustainable engineer is able to capture at any one time. The engineer, rather than trying to fight this complexity, must learn to work with it appropriately.

Finally, much complexity, and certainly a lot of the institutional failure to deal well with complexity, arises from the rapid and accelerating evolution of technology systems, especially ICT. This has very important implications for sustainable engineering, because many of the available tools, approaches, and frameworks of relevant fields such as industrial ecology and lifecycle assessment traditionally focus on material and energy flows. This is obviously a useful focus but, because it leads many, especially in the sustainability field, to overlook the complexity of the systems emerging in the real world and the growing role of ICT and cyberspace in culture, it is increasingly inadequate. Information and cultural structures, especially those associated with radical technological evolution, cannot continue to be ignored and must be regarded as critical domains of sustainable engineering, even if the means to explore them rigorously and our basic understanding of them is still fairly primitive. It is not, therefore, that sustainable engineering studies and practice should not focus on material and energy flows; rather, it is that the field as a whole must learn to treat cultural and informational realms with the same sophistication and seriousness as physical realms.

3.5 MODELS, IDEOLOGY, AND COMPLEXITY

Complex systems cannot be understood directly, and yet they are the essence of sustainable engineering. Accordingly, complicated computer-based models are one of the most important tools of engineering in general, and sustainable

engineering in particular, as well as underlying disciplines such as environmental science. And yet more and more scientists are arguing that quantitative modeling, the basis of much environmental science, cannot predict outcomes of complex natural processes. This obviously raises a challenging and very fundamental issue: How do we as engineers and technologists design and manage in a world where the complexity of the relevant systems—the Everglades, the climate system, the nitrogen cycle, urban systems—assures that any coherent framework such as a model with which we approach the issue will, by virtue of being coherent, necessarily be incomplete and inaccurate? Unpacking this dilemma may help us use models, which, after all, are the only real tool we have to try to understand and address highly complex systems, rationally and ethically.

Let's begin by thinking about what a model is and how we build one. Because we only use models where we can't understand the system directly, any model begins with the intelligent simplification of a more complex reality by capturing the information that is "important" and "relevant," and ignoring "nonrelevant" information. In other words, we adopt a particular worldview or set of complementary rules that tell us when something is "relevant," and thus to be included in our model, and when it is not. If something isn't relevant, we can, for purposes of the model, ignore it and all its supporting data. This is a perfectly legitimate way to do engineering and science. Newton, after all, developed his laws of motion by ignoring minor forces (such as friction) that would have made his simple, and therefore powerful, conclusions far more complicated and confusing. How do we know what information to drop when constructing a model? As with any observation of a complex system, that is a function of the reason for which the model is being created, the structure of the underlying system, and the ability of the modeler. It is neither pure creation of the observer, nor is it a product of the underlying system; it partakes of both.

A properly created model is therefore one that embodies the information necessary to understand the query posed by the modeler, and only that information, because excess unnecessary information only adds complexity that, by definition of the model, is unimportant to what we're trying to understand. Each model will then have boundaries within which it is useful—that is, within which the information and model structure are appropriate to the queries addressed to it. But the boundaries are frequently not obvious; especially in highly politicized debates, there will be powerful tendencies to extend the model beyond them. Thus, for example, models of global climate change create probabilistic future scenarios that can inform public debate; to present those results as inevitable, or demanding certain social or policy responses, however, is invalid. To begin with, they include primarily scientific data, because they are put together by communities that value such data as a valid way of understanding the world. But even then, they are profoundly incomplete.

How so? Among other things, any model projected into the future makes assumptions about technology states that are almost inevitably wrong. We cannot know what technologies may exist fifty years from now, but we should at least be honest enough to admit that, and to understand models that make assumptions about technology states are useful thought experiments, but no more. In the case of climate change, for example, we know that technological changes will create dramatic differences in human relationships with the carbon (nitrogen, hydrologic, and other) cycle over the next fifty to one hundred years, yet the models project smooth curves for over a century. This is clearly wrong—we may not know how, but we know it's wrong. But few scientists are students of technology, so they focus their models on what they know, which is the science. Again, this is not wrong, but it is a limitation that few seem to understand, and fewer still acknowledge. More subtly, models are inevitably part of highly reflexive systems, which means they feed back quickly into the system (in this case, societies, economies, and cultures) that they are purporting to model. Climate models, by existing, change behavior and institutions in ways that invalidate their initial assumptions. Thus, to the extent climate models are right, they make themselves wrong: All good models undermine themselves over time. Yet, it is clear that we need models to explore the complexity of the world we are creating. What is to be done?

In answering this question, we should differentiate between problems that can be solved with **reductionist** approaches and those that necessarily involve complex systems, where behavior emerges from the system as a whole. *Reductionist* simply means that you can understand and manage the system by breaking it down into smaller and smaller pieces for analysis, by "reducing" it. Thus, for example, if your job is to engineer a fruit salad, you can do that by breaking it down into its constituents—apples, melon, oranges, pineapple—and cutting them up individually. But if your job is to design and manage the Everglades, or the climate system, you cannot do that. The problems you're concerned with, such as the relationship between agriculture, water flow, invasive species, and biodiversity in the Everglades, emerge from the system as a whole, not just from its parts. The tools of reductionist science and engineering are familiar ones: direct observation or replicable experiments. Understanding the emergent characteristics of complex systems, however, requires increasingly complicated models.

So, models are both useful and necessary in a complex world. Indeed, the difficulties with models arise primarily from their misuse in discourse, rather than from the models themselves. That models are partial reflections of complex systems and, change policies and management options in unpredictable ways does not mean they are wrong in themselves. But the ease with which quantitative model results can be mischaracterized, and the strong incentives to do so by those involved in contentious discourses, means that models in our

society are highly susceptible to abuse. Moreover, the tendency to overstate and misuse the predictive and analytical capabilities of models, by both scientists and activists, has serious implications, for it undermines not just the particular models involved, but the image of objectivity that is itself the strongest validation of the scientific and technological discourses.

Consider, for example, climate change modeling. No one has ever observed climate change directly—indeed, no one ever could. Rather, we understand its existence from first principles (physics and atmospheric chemistry, as discussed in Chapter 1) and a series of observations that, taken together and integrated in models, lead to the assertion that if the models are correct, anthropogenic global climate change exists and will have certain effects. The constellation of data, models, and projections are then simplified into an assertion of reality in public dialog, to the point where some governments, such as the United Kingdom, are considering issuing each citizen individual emissions permits, and many climate scientists and activists are demanding immediate cuts in consumption and fossil fuel use in developed and developing countries alike. The sense of reality, indeed catastrophe, derived from model predictions, is powerful. Our problem is that once we bury an assumption in a model and create numbers around it, we believe it is real.

There is at this point little question that poorly understood anthropogenic changes in many natural systems, including climate, are occurring. Nor is there any question of the need for models as tools to explore potential systems behaviors, especially undesirable ones. The problems arise when people speak of model projections and scenarios in language that reifies them: Predicted patterns "will" happen, or certain phenomena "are the result of climate change." This language is used partly because mere scenarios lack the authority to drive the social engineering that climate change activists perceive as necessary and desirable. If you are going to be heard above the cacophony of 24/7 news channels, the Internet, Facebook, Twitter, and the busyness of the world, you need to have a certainty and a looming catastrophe that models cannot provide. The risk, of course, is that at some point the members of the public may become aware that they have been stampeded by deliberate manipulation of information and turn against both the environmental community, and the scientific and technological community more generally, with the unfortunate prospect of undercutting efforts to address real and serious problems.

But it is not just models that the sustainable engineer needs to be careful with. After all, in some ways models are just quantifications of underlying worldviews and ideologies, and these also must be understood by the modern technologist. Again, it is not that belief systems or worldviews are inherently bad; we all have them, and they are a necessary part of the way we deal with a highly complex and difficult world. There is, however, little question that ideology and rigid worldviews are particularly problematic at the dawn of the

anthropogenic world, which as we have seen is characterized by exceedingly rapid, unpredictable, and profound change in fundamental relationships and systems involving natural, built, and human components of extraordinary static and dynamic complexity. Like models, any successful ideology, from marxism to free market capitalism to environmentalism, is necessarily a simplification of reality; in fact, that's usually an important part of its mass appeal. But because the elements and structure of this simplification necessarily lie in the past, not the future, they embed assumptions and implications that are necessarily increasingly anachronistic in a period of rapid and discontinuous change. Ideologies, then, end up trying to apply past truths in a relatively simple way to a world changing and growing complex by leaps and bounds. That's a mismatch.

More subtly, ideologies that become dominant block information transfer and dialog, and thus inhibit innovation and technological evolution—again, an undesirable effect in a rapidly changing world. There is much evidence that cultural and technological innovation and evolution occur in societies and institutions that are not dominated by any single perspective, but that allow, even encourage, managed conflict.

The solution is not for the sustainable engineer to try to rid himself or herself of all belief systems—that's probably impossible, anyway. The need is for sustainable engineers to understand the power of ideology, recognize it in one's own beliefs, and be able to respect and work with different belief systems in the design and management of complex problems without trying to impose one's own worldviews on the result.

Exercises

3.1 Many types of engineering problems depend on understanding issues of scale and boundary definition, and there may be multiple scales involved in understanding the operation of a particular system. What boundary of the "city of Phoenix" is implied by:

a. Air freight transportation?
b. Financial networks?
c. Water delivery systems?
d. Sewage systems?
e. Immigration?

3.2 You are in charge of implementing a large-scale program in Africa to extract methanol for transportation from a root tuber that is widely grown. You have the mandate to consider a variety of factors, including social, environmental, technological and economic, in your preliminary design. How would you consider gathering all the necessary information you need for your design and how would you put it all together? Where do you expect to encounter show-stoppers, if any?

3.3 The rapid rate of change in modern society, especially the accelerating evolution of technology systems, would argue for an adaptive, flexible politics and institutional style.

Instead, one seems to find more argument and resistance to change. Using your country as an example, do you think this is the case, and if so, why? And is there anything that can, and should, be done about it? What role should engineers play in these public discussions?

Annotated Bibliography

1. S. Kauffman, 1995, *At Home in the Universe* (Oxford: Oxford University Press), is a good nonquantitative introduction to the subject of complexity. There are many other useful books on complexity, of course, many of which increasingly focus on network theory. A good introduction to that area would be A. Barabasi, 2002, *Linked: The New Science of Networks* (Cambridge, MA: Perseus Publishing). In addition, many engineering and scientific journal articles cover complexity in various aspects: biological and ecological systems, social networks, engineered systems, and the like. The challenge is not finding reading material on complexity, but finding reading material that provides useful and relevant insight.
2. Those interested in a readable introduction to complexity in business management and industrial organization, as well as a description of the "beer game" mentioned in this chapter, might want to start with P. Senge, 1990, *The Fifth Discipline: The Art and Practice of the Learning Organization* (New York: Doubleday). Management books on complexity are useful for sustainable engineering because, like engineers, managers have to grapple with the world the way it really is, not as they might wish it to be, and thus have to be pragmatic about what public values really are, instead of assuming that everyone agrees with them.
3. Those who are inclined to trust models as real predictions of the future, as opposed to using them as tools to create and explore scenarios to help understand possible future challenges, might want to read O. Pilkey and L. Pilkey-Jarvis, 2007, *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future* (New York: Columbia University Press). The point is not that models are "bad," just that they are easy to misuse, especially by ideologically committed activists, including scientists.

CHAPTER

4 Sustainability

4.1 INTRODUCTION

Early in his career, the author of this book was assigned responsibility for implementing industrial ecology and design for environment for a large multinational firm. This anonymous firm had an excellent research and development laboratory, Bell Labs, and at one point the author went to a design team for a telephone, suggesting that the firm design and market a "sustainable telephone." There was widespread enthusiasm until one of the design engineers asked him what a "sustainable telephone" actually was. It is best not to go into painful detail regarding the subsequent discussion, except to note that the author was completely unable to provide an answer, which was not a good thing. Oh, eventually we worked out some heuristics—some rules of thumb—that could be applied, such as "all things equal, use less toxics in design and manufacture," but it is not an experience you should wish to replicate. Time has passed, and you need to be able to do better.

These days, of course, many students have heard of sustainability, and most feel they have a pretty good idea of what it means. Yet, when classes are surveyed, one tends to find that students either equate sustainability with environmental protection, or with increasing efficiency, especially as regards energy and material use. It arguably includes both, and yet it is neither. For an engineer challenged to practice in our rapidly changing world, it is important to understand the ambiguity and fuzziness of the idea of sustainability, and at the same time to realize that the public, employers, and stakeholders are demanding "sustainable engineering." This chapter discusses the basics of sustainability,

how that term interacts (and to some extent conflicts) with the theory and practice of engineering, and how the sustainable engineer can—indeed, must in this modern world—serve both sustainability and engineering.

4.2 SUSTAINABLE DEVELOPMENT

Although there had been considerable discussion regarding the interplay between economic development, especially in less affluent countries, and global environmental quality since the environment became a politically potent issue in the late 1960s, the definitive definition of sustainable development was not provided until 1987. That definition was provided by the World Commission on Environment and Development (WCED), established as an independent commission by the United Nations in 1983, in its report, *Our Common Future*. Since then, there have been literally hundreds of different interpretations of *sustainability*, with an accompanying industry in producing sustainability metrics that range broadly across environmental and social policy. This profusion of riches has had the effect of making "sustainability" a difficult concept for those who have not spent years in this area; indeed, many have (a) given up on the idea entirely; (b) used the confusion and ambiguity surrounding sustainability discussions to dismiss the issues behind sustainability entirely; or (c) become members of an introverted, insular, and occasionally sanctimonious sustainability community. For us, however, this ambiguity need not be an insurmountable barrier.

Because the WCED's statement of the theory of sustainable development was the first and remains the most authoritative, let's begin there (*sustainable development* and *sustainability* will be used interchangeably in this discussion). The Commission began by defining *sustainable development* as development that "meets the needs of the present without compromising the ability of future generations to meet their own needs."[1] It "requires meeting the basic needs of all and extending to all the opportunity to fulfill their aspirations for a better life," assuring that the "poor get their fair share of the resources. . . ." The Commission saw global inequity and poverty as a primary cause of environmental degradation: "many problems of resource depletion and environmental stress arise from disparities in economic and political power" (p. 46). For the Commission, this required that "those who are more affluent adopt life-styles within the planet's ecological means" and a strong "concern for social equity between generations, a concern that must logically be extended to equity within each generation" (pp. 9 and 43).

[1]The World Commission on Environment and Development, *Our Common Future* (Oxford: Oxford University Press, 1987), p. 8.

4.2.1 Pillars of Sustainable Development

In understanding sustainability, it is important to begin by realizing that sustainability is a classic example of a cultural construct: a symbol, idea, or phrase by which societies create and transmit meaning. Cultural constructs sometimes arise on their own, but frequently they are deliberately invented to meet changing circumstances or to advance specific ideas about the world. In the case of sustainability, of course, the invention process is fairly explicit: The concept was initially popularized by the book mentioned earlier, *Our Common Future.* Moreover, it is not simply chance that the Commission that created that report was a child of the United Nations, which has interests in both economic development, especially of poorer areas of the globe, and environmental quality. That's because a major challenge for the international community in the 1970s and 1980s was to reduce conflict between two important discourses: the economic development discourse that sought to encourage economic growth, especially in developing countries, and the environmental discourse that sought to preserve as much biodiversity and unspoiled land, especially tropical forests, as possible. The solution, at least conceptually, to this conflict was to develop the idea of sustainability, which combined both environmental protection and development in one framework.

Viewed in terms of invented cultural constructs, sustainability is a runaway success. In less than twenty years, the idea of sustainability evolved from an explicit construction derived from a particular political philosophy, to adoption as an integral part of modern environmentalism, to becoming a major policy discourse in itself. But a few important caveats should be noted. First, simply because someone can invent an inclusive cultural construct does not mean that underlying realities have shifted. That is, just because we now have the popular *idea* of sustainability does not mean that we can *in fact* have both environmental quality and economic development. After all, just because we as humans can invent the idea of unicorns does not mean the world is suddenly full of unicorns; they remain imaginary. In addition, we cannot expect sustainability to resolve particular conflicts between social and environmental goods that may arise in, say, a particular design project; difficult tradeoffs and balances will depend on the facts and preferences in specific cases, not on more rhetorical references to grand visions of sustainability.

With this in mind, it is useful to explore the two primary themes of sustainability: environmental quality and egalitarianism. Let's look at them in turn.

For most people, if sustainability stands for anything, it stands for environmental values. Thus, for example, even many engineers working in the field tend to equate "green engineering" with "sustainable engineering," although the latter is clearly a broader and more complex domain in that it includes social and cultural dimensions as well. But sustainability tends to favor a particular approach to environmentalism, a human-oriented approach. Figure 4.1 shows this in

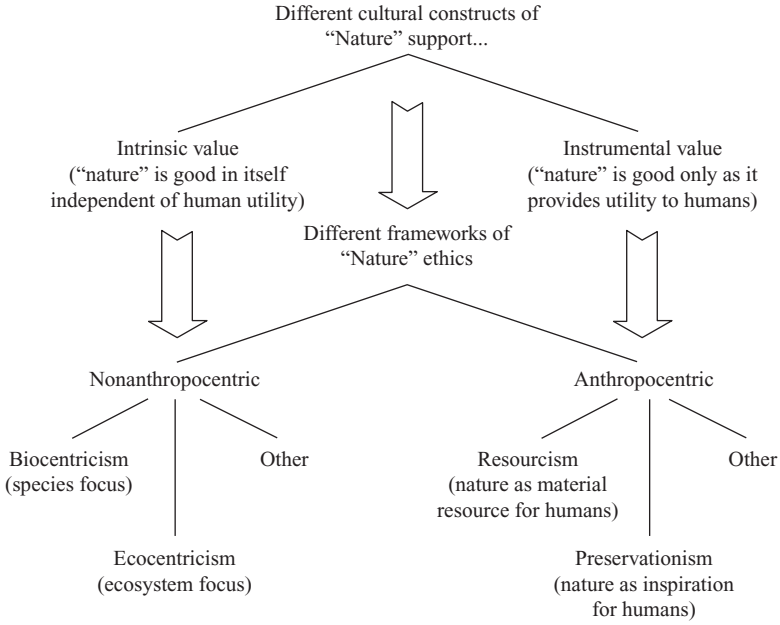


FIGURE 4.1 Valuing Nature. This diagram shows the split between those who value nature in itself and those who value it for what it can provide to humans, and some of the ethical postures that flow from this split. In practice, these clear delineations often become confused: in the United States, for example, the National Environmental Policy Act balances development against environmental impacts, a view compatible with sustainability, while the Endangered Species Act allows no such balance, very much tracking the "intrinsic value of nature" worldview.

(Based on: Proctor, J. D. "Whose Nature: The Contested Moral Terrain of Ancient Forests," in *Uncommon Ground*, ed. W. Cronon. New York: W. W. Norton & Company, 1996, pp. 269–297.)

more detail: Either we can value "nature" (a very complicated cultural construct in its own right) for what it can provide to humans, or we can value "nature" for itself, as something that demands respect independent of any human value or purpose, and that has rights at least the equal to any human rights. The value systems that flow from these different approaches are quite different. Those who value nature instrumentally tend to discuss the monetary benefits of natural systems: Rare ecosystems support tourism, swamps and salt marshes provide services such as water storage and treatment, rain forests store greenhouse gases, managed forests provide a stream of renewable resources for industrial purposes. Those who value nature for its own sake tend to reject the idea of putting a monetary value on any natural system, and will focus on the independent value of a particular organism, species, or ecosystem. In the extreme, this position can come to view nature as opposed to people, and result in physical attacks on people and property in the name of "nature." Sustainability, as a cultural construct derived in part from the economic development discourse, takes an instrumental

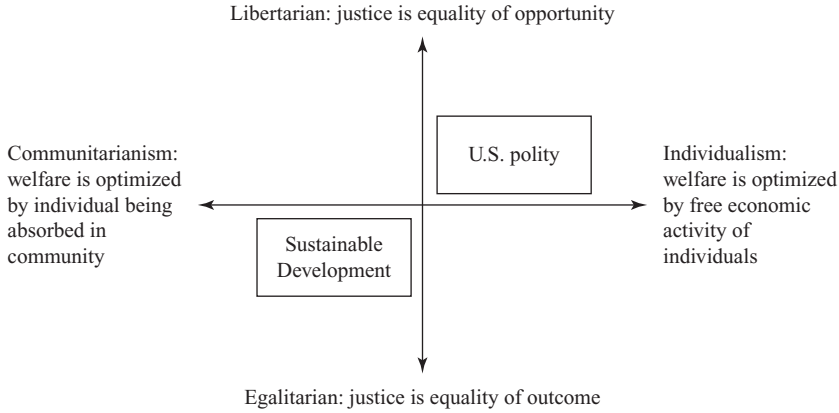


FIGURE 4.2 Sustainability and Basic Political Values. This figure shows that sustainability tends to favor the communitarian and egalitarian political perspectives, a political philosophy quite popular in Europe, as opposed to the libertarian and individualistic political perspectives more characteristic of the United States. The result is a higher level of skepticism regarding the idea of sustainability in the United States and, partially to overcome this, a much stronger emphasis in the United States on sustainability as an environmental, as opposed to social, concept.

view of nature, and thus tends to be rejected by those who see intrinsic value in nature. This is why, for example, many "deep greens"—ideologically committed environmentalists—are skeptical of sustainability, and why the sustainable engineer should not assume all environmentalists take the same view of sustainability and its implications for engineered systems.

Similarly, sustainability takes a clear perspective on issues of social equity, or how power, wealth, and income should be distributed in a "just" society. This is, of course, also a very complicated area, but one can construct a fairly simple compass diagram that explains the basic framework of the discussion. Figure 4.2 has two axes: One extends from those who view welfare as optimized when the individual is subordinate to the social community, to those who view the individual as supremely important. The second axis extends from those that view justice as requiring equality of resources between individuals, so that wealth and income are spread among everyone relatively equally (the "egalitarian" perspective), as compared to those who believe in equality of opportunity, which necessarily leads to inequality of outcome, as those who study hard, work hard, and succeed do better than those who choose for whatever reason not to (a "libertarian" viewpoint). For the latter, redistribution of income is itself immoral, because you must take from those who earned more through their individual efforts and give it to those who didn't, and the only way to do that in a society is through implied force. In practice, of course, most societies combine elements of all four extreme positions, but this compass does provide a striking idea of why the cultural construct of sustainability has been more successful in some cultures than in others.

This mapping of sustainability makes a few important points clear. First, the reason why sustainability has always had a more difficult time catching hold as a strong concept in the United States is apparent: it embeds social values that are comfortable for Europeans with their history of social democratic parties, but that are much less mainstream in the United States, which tends to remain more individualistic and libertarian than most other countries generally, and than European countries specifically. Box 4.1, for example, discusses a major U.S. presidential sustainability initiative that, in fact, virtually disappeared from sight, in part because those involved failed to understand the implications of Figure 4.2. Second, there are no "right" answers to these sorts of values questions. Each of us will have a personal perspective on the degree to which communitarianism or individualism is better, or whether more redistribution of income is "just" or not. For a problem solver, which engineers are, the real question is not what your personal views are, but what views are held by which stakeholders and why, and how able they are to compromise to achieve some sort of consensus design. In other words, you should use Figures 4.1 and 4.2 to understand your stakeholders, and the context within which your particular project or system is embedded, so that you can craft more workable solutions.

4.2.2 Sustainability as Myth

What is sustainability? This question actually can be broken down into two separate questions: First, what is the current operational definition of *sustainability*? And second, what cultural role does the concept of sustainability play? Sustainable engineering requires an understanding of both these questions.

The obvious starting point for any operational definition is to identify what it is that we are trying to sustain. Despite much rhetoric to the contrary, sustainability is not about "saving the planet," for the simple reason that the planet is not going to disappear. Some aspects of it may reflect human occupation—its emissions spectrum, for example, as television reruns and light from all our cities routinely escape into space—but the planet itself will continue to evolve with or without humans. It is not "life," either, because our ability to threaten the existence of life is virtually nonexistent; we can't even cleanse our hospitals of bacteria (although clearly the evolutionary paths that life takes are affected by human activity). Nor is it in all likelihood "humanity," because humans are great generalists, so it is likely that if any megafauna continue to exist on the planet, we'll be among them. But it may be true that the stability of current social, cultural, economic, and demographic structures are challenged by the evolving anthropogenic world. This raises the ironic scenario that, despite all the emotion, the only real contribution of the sustainability dialog is to preserve existing power structures and social hierarchies, a highly conservative, indeed reactionary, result.

BOX 4.1 The President's Council on Sustainable Development

Few people today remember, or have heard of, the major sustainable development policy initiative of the Clinton/Gore administration in the United States. By Executive Order No. 12852, dated June 29, 1993, President Clinton established the "President's Council on Sustainable Development" (PCSD). The general duty of the PCSD was to advise the President on matters involving sustainable development. More specifically, it was asked to develop and recommend to the President a national sustainable development action strategy that would foster economic vitality and help the President conduct an appropriate awareness and participation campaign. The PCSD was cochaired by Ray C. Anderson, the chairman, president, and chief executive officer (CEO) of the flooring company Interface, Inc., and a well-known corporate sustainability campaigner, and Jonathan Lash, the president of the environmental NGO, World Resources Institute. The members of the PCSD included numerous environmental, industrial, and governmental luminaries. It was in existence from 1993 to 1999; over that period the PCSD published many reports and papers, and held numerous public meetings around the country. It ended with a spectacular bang, the "National Town Meeting for a Sustainable America," on May 2–5, 1999, at the Convention Center in Detroit, Michigan, which the organizers firmly believed would "catalyze a national movement toward sustainable development, with events in Detroit, Michigan and Points Across America." (More information can be found at clinton2.nara.gov/PCSD/, accessed August 2009.) The PCSD optimistically explained its rationale for such a meeting:

> **Why a National Town Meeting?**
> Momentum is building for the notion that America can prosper economically without sacrificing the environment and quality of life. The National Town Meeting will highlight the work of communities, businesses and individuals who are offering new solutions to help America grow in a sustainable way. The NTM will highlight the work of these Americans and will offer solutions to America's sustainability challenges.

Thousands attended the Town Meeting, and thousands more participated in the satellite events around the country. In 1999, all appearances were that the PCSD was a smashing success.

Yet, despite all this activity, a six-year span of articles, publications, pamphlets, and exhortations, participation by major industrial and environmental institutions, and an administration committed to environmental values, the Council today is virtually unknown, and few if any of its recommendations have been implemented, nor has anything come of its momentum in 1999. Why? The answer is of some importance for sustainable engineering because in some ways—identifying an important problem and trying to solve it in the real world—the PCSD was trying to do what sustainable engineers must do all the time.

It is of course impossible to answer the question definitively, but several observations may be worthwhile. First, although the Council covered the right institutional

bases, it appeared to at least some outside observers that it did not successfully include those who might have been skeptical of its activist stance.[2] Inclusiveness was valued as an abstract concept, but never achieved, and perhaps not seriously sought. The result was that it frequently seemed that the Council members were talking to each other and to the "true believers" they attracted to their meetings, rather than trying to achieve serious political synthesis across boundaries. Despite appearances of openness, it became a closed dialog. In part this may have been because it was not clear, at least to some, whether the purpose of the PCSD was to explore the topic of sustainability in an American setting or to market sustainability to Americans—two very different tasks. For the sustainable engineer, the lesson here is to be very clear on agendas and real goals, and that means the hidden ones as well as the formal, stated ones. The confusion regarding agendas is also still somewhat characteristic of the sustainability discourse, which is something the engineer and technologist should bear in mind when working with sustainability activists.

More important, the PCSD never grappled with the issue raised by Figure 4.2 in the text. General homilies about sustainability are effective in Europe because the underlying framework is not problematic in the European context: egalitarianism is a priority value, especially in northern Europe and the European Union. But homilies and generalized statements of goodwill and social value are inadequate in a cultural context where conflict exists over the basic political framing of the cultural construct itself—where, in other words, there is no agreement that egalitarianism is more important than libertarianism. This cultural discontinuity was never adequately recognized nor addressed by the PCSD, and is undoubtedly an important reason for its virtually complete failure.

For the sustainable engineer, perhaps another very important lesson is that, regardless of how we might wish the world, successful problem solving in the real world requires accepting the world as it is. Perhaps in an ideal world it would not be necessary to include contrarians and those who disagree with you in a dialog such as the PCSD carried out, but the world is not ideal. Thus, the sustainable engineer should take great care to obtain all relevant input and not just the input that supports a favored design decision or project direction. In the short term, such unanimity might result in apparent success, as the Town Meeting suggested. In the long term, however, solutions must be economically, politically, and culturally viable, which means that engagement with the diverse opinions and complexity of the real world is unavoidable.

Despite the incoherence of the sustainability discussions, there is another view of sustainability that is less cynical and perhaps more interesting. Ask yourself what has changed over the past twenty years, as sustainability has become

[2]Full disclosure: the author was asked to serve on several subcommittees of the PCSD and declined for a number of reasons, although he continued to be engaged with the process on an informal basis. These observations, therefore, are neither based on personal "insider" knowledge, nor are they the result of formal study or analysis; rather, they reflect experience at the time and common sense. They should therefore be taken with the appropriate grain of salt.

such a topic of discussion. The answer is obvious: by almost any measure, in virtually every society, the world has become more complex. Anxiety levels over this complexity are very high, yet the vast majority of people have neither the time nor the expertise to engage with any aspect of it, let alone the anthropogenic Earth as a whole. At first blush this situation appears to the engineer to display serious cognitive dissonance: We can't tell you what sustainability is or means, but you had better learn to be competent in sustainable engineering. This very strong social demand should not be dismissed, nor ignored, because at heart it reflects a real public concern about a world that really has gotten more complex—and it is thus part of the engineer's design environment. Your customers demand sustainable engineering. So what is to be done?

The first thing is to recognize that designing, building, or manufacturing a "sustainable anything" is essentially impossible, because, despite hundreds of different definitions and thousands of proposed metrics, no one really knows what a "sustainable anything" actually is. Sustainability is more usefully understood as an evolving social myth, where *myth* is not used as a pejorative term, but in the sense of a foundational narrative that helps individuals make sense of a complex and unpredictable world. It is a way of combining values, beliefs, anxieties, and fears into a framework that helps people feel in control and derive meaning. Myths are not analytical and instrumental, just as "sustainability" provides no guidance to an engineer, and they are not quantifiable. Yet they are emphatically not meaningless for that reason, because as they gain power they structure the way people think about and react to their world—indeed, how they perceive it. Once sustainability is understood as an evolving myth enabling modern humans to deal with an increasingly complex world, the language and culture that surrounds it becomes more intelligible. Comments that "the planet" is "suffering" make no sense to a scientist, because planets cannot suffer; but in a mythic structure, suffering of nature, the individual, and society are frequently coupled. Moreover, sustainability as myth explains the sometimes apocalyptic language that surrounds sustainability and particular issues such as climate change. We often hear, for example, even in the scientific literature, statements to the effect that "we're the last generation on Earth that can save the planet." Of course, this sort of assertion is neither simple truth nor defensible in any factual context, but the idea that the end is near, and that only immediate, dramatic action can avoid it, is typical of mythic (indeed, religious) structures.

But, we might argue, sustainability is a scientific concept; just listen to the language that people use when they talk about it—the climate scientists, the conservation biologists. And it is certainly true that the language of sustainability is based on the scientific and technological discourses, which are the recognized sources of secular truth in Enlightenment cultures. In this way, sustainability arguments reflect contemporary understanding and values, and can

be coupled to the authoritative language of science. This power is needed, of course, for sustainability to be an effective modern myth. The cultural construct of sustainability, however, is an objective façade over a subjective structure; unless understood as such, it can significantly mislead those who attempt to apply what they conceive as an objective set of criteria to technological or engineering situations. In fact, in some ways it is the failure of sustainable engineering to understand "sustainability" as a guiding myth, rather than a source of quantified design objectives and constraints, that perhaps causes the major conceptual problems regarding sustainable engineering.

It is easy to be dismissive of sustainability, and many sustainability advocates unfortunately display attitudes that may encourage that response, especially among those who are technically trained. It is certainly the case that sustainability is still a young and inchoate myth. But it is important to many people, and it is based on changes in the world that, absent some sort of collapse, are leading to unprecedented challenges to society and technocrats alike. The reality of the belief and of the changes underlying it cannot be responsibly ignored by any engineer in today's world. Sustainability as myth makes sustainable engineering more, not less, important to engineering professionals, to their clients, and to society as a whole.

4.3 SUSTAINABILITY AND RESOURCES

Material selection is a common activity in virtually any branch of engineering, either directly, as in product or infrastructure design, or indirectly, by choosing technologies that in turn require particular material inputs. The most obvious example of the latter, perhaps, is the myriads of technologies that by requiring energy, imply substantial demand for fossil fuels across the economy as a whole. Understanding a little about materials and how they function is therefore a critical aspect not just of the sustainability dialogs but also of sustainable engineering.

Although there are many ways to differentiate materials, four major categories can be identified based on shared characteristics: metals, energy materials, industrial nonmetallic materials, and construction materials. Metals tend to occur in fairly low concentrations in the Earth's crust, especially because high-concentration ore bodies have been exploited for centuries. Fossil fuels are the dominant form of energy materials, although nuclear materials on the one hand, and renewables such as wood and brush on the other, are also part of this category. Industrial materials might include different types of clay and silicates, salt, and other products. Finally, construction materials dominate in terms of volume, but not in terms of value per unit weight. Each category implies certain infrastructures: construction materials, for example, are both heavy and low value, and are thus not shipped long distances.

	Activity	Soil, Rock and Silt Moved (billion tons per year)
Nonhuman		102.5
	Wind erosion	1
	Glaciers	4.3
	Mountain building: Continental	14
	Mountain building: Oceanic	30
	Water: Long-distance sediment transfer	14
	Water: River meandering	39
	Water: Erosion from ocean waves	0.24
Human		40–45
	Building roads	12–13.8
	Housing starts	3.2–3.7
	Mining	15.2–17.5
	Agriculture and other activities[1]	10

[1]Contribution of human disturbances such as those associated with agricultural activities to the long-distance sediment load of rivers.

FIGURE 4.3 Estimated Average Annual Soil and Rock Transport (1991 Data). This data indicates that humans, responsible for roughly 40 percent of the overall material transport, have become one of the largest "earth systems" on the planet. Data is the latest available; if anything, the human impact has become more pronounced since 1991.

(Data from: Hooke, R. L. "On the Efficacy of Humans as Geomorphic Agents," GSA Today 4, no. 9 (1994): 217–225.)

Any discussion regarding human mobilization of materials should recognize several patterns. The first is that humans have grown to dominate many material flows, at least in terms of critical dynamics. Chapter 1 discussed this in regard to carbon and nitrogen cycles; the same is true of many different materials, as Table 1.5 indicated for a number of elements such as arsenic, copper, nickel, and lead. Equally important, the volumes of material that humans now manipulate have grown to a scale equivalent to major planetary-shaping forces, such as erosion (Figure 4.3). Although the preponderance of this transport involves construction materials, it nonetheless is yet another demonstration of the complex integrated human/natural/built systems that characterize the Anthropocene: activity associated with bulk materials, for example, has obvious impacts not just on physical systems, but on ecosystems and more generally biodiversity.

Scale is important for sustainable engineering from another perspective as well. Much design occurs at a small scale: the prototype unit, bench tests of particular chemical treatments, and the like. Accordingly, questions of material availability seldom arise in this stage of research and development. Yet in a global economy, the scale at which a technology may be deployed can have

significant environmental, social, or production implications: indium, for example, may be readily available in relatively small volumes, but to suddenly substitute indium solder for tin–lead formulations throughout electronics would seriously distort production and supply infrastructure, at least until **resources** could be converted to **reserves** (see Box 4.2 for the difference between these two terms). A good example of the failure to consider the implications of scale involves arsenic use in the United States. Historically, arsenic, primarily a by-product of copper production, was used as a pesticide and desiccant. Recognizing that it was inappropriate to use a material with serious health implications in such a dissipative manner, Environmental Protection Agency (EPA) regulations in 1987 and 1993 essentially phased out this use. However, the level of U.S. arsenic use remained relatively constant, as use shifted to wood preservation, primarily as chromated copper arsenate (CCA). Approximately 48 billion board feet or 113,000,000 cubic meters of wood were pressure treated with CCA between 1985 and 1999; U.S. consumption of pressure-treated wood accounted for 90 percent of worldwide arsenic demand at the end of that period, and made the United States the world's largest consumer of arsenic. Note that this also is a dissipative use. It is not that CAA treatments weren't effective, or that it may not have been appropriate in some limited uses where conditions and cost/benefit considerations were favorable. But this, of course, is not what happened, and the scale up of a successful bench technology had the systemic result of substituting one dissipative use of a toxic material with another. Technology system scale, in other words, must be part of material choice, and product and process design, from the beginning.

Perhaps the major question raised by most people looking at data on human material use is deceptively simple: Is there enough? The anxiety associated with this question is heightened by two factors: First, it is obvious that at some point by definition finite resources "run out"; and second, there are a lot of alarmist articles, especially in the sustainability literature, to the effect that we are running out sooner rather than later.

We can begin thinking about this issue constructively by making a fundamental point. There are clear limits if one looks at **reserves**—materials recoverable using current technology and assuming current economic and regulatory conditions. If, however, one considers **resources**—the amount potentially recoverable—the uncertainty about how much is available is much greater, but there are few apparent absolute limits in the short term (see Box 4.2). Thus, if one wishes to be alarmist, one would divide "proven reserves" by per capita use and come up with a prediction of imminent shortfall; for the engineer or analyst, however, the more important question is usually one of absolute scarcity—resources per capita, a much fuzzier but much larger number.

Thus, for example, one of the areas in which scarcity is much claimed these days is fossil fuels, especially the "peak oil" idea that petroleum production

BOX 4.2 Reserves and Resources

In discussing the availability of minerals, it is very important to understand the difference between various ways in which the potential stock of a particular mineral are expressed. If, for example, someone takes the "proven reserves" of a particular metal and divides it by per capita demand, that person often gets very frightening predictions: "We will run out of indium in five years!" This is at best uninformed and misleading, and at worst deliberate scaremongering.

Consider the categories in Figure B4.2.1, which present the general categories of mineral stocks (including fossil fuel stocks). Proven reserves is the most restrictive category and is defined as the subsurface quantity of a particular material that is (a) known and (b) profitable to exploit given existing technologies, prices, and other relevant conditions. For oil companies, for example, proven reserves must be identified by exploratory well, be commercially viable, be in ground, and be recoverable using existing industry technology. In the Unites States, proven reserves are the only type that the Securities and Exchange Commission allows to be reported to investors, which means that companies are very conservative in such estimations (because to report reserves as proven that may not be opens up a company to civil and potentially criminal litigation). The reserves base (or "unproved reserves") consists of reserves that meet most of the requirements for proven reserves, but because of technical or legal uncertainty (and corporate conservatism) may not be appropriately

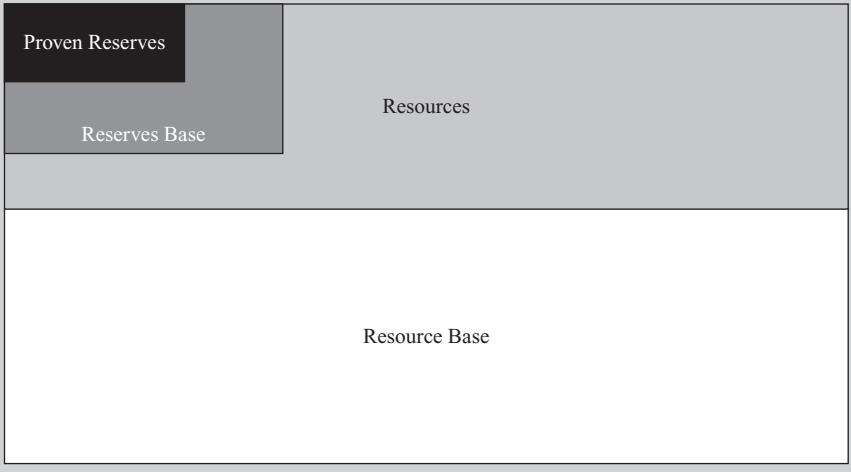


FIGURE B4.2.1 Mineral Resources (Expressed in Order of Geological Significance). This figure provides a schematic illustration of the four major categories of mineral stocks. Analysis of mineral availability for sustainability purposes that rely only on "proven reserves" can be unduly alarmist and suggest illusory shortages of resources.

(Based on: Information presented by Andrew Bloodworth, Head of the British Geological Survey, Workshop on Sustainable Materials, Lausanne, Switzerland, July 2009.)

categorized as proven reserves. Reserves as a general category are important because they are "economically real"—that is, once a reserve is established, it can be used by firms as security for loans and the like.

"Resources" constitute a much larger category, and consist of mineral concentrations where the location, grade, quality, and quantity are known or can be estimated with reasonable certainty based on geologic analysis and evidence. This category is much larger than most people realize, but difficult to quantify in part because of the greater degree of uncertainty compared to reserves, and in part because many companies don't bother to narrow down the value of their resources until economic, political, and strategic considerations make production from a particular location possible. Moreover, those who are not mining engineers may assume that more is known about mineral formations than is the case. One example is the iron oxide/copper/gold (IOGC) deposit model, which is fairly recent. The Olympic Dam IOGC in Australia, for example, discovered only in 1975, is the fourth largest remaining copper deposit, the fifth largest gold deposit, and the largest known uranium deposit in the world. New discoveries are announced with fair regularity. Consider Oyu Tolgoi in Mongolia, with an estimated 70 billion pounds of copper and 32 million ounces of gold, announced in March 2007, or Reko Diq in Pakistan, with 47 billions pounds of copper and 38 million ounces of gold, announced in March 2008.

Finally, there is the **resource base**, which simply consists of all of a mineral commodity contained in the Earth's crust. In all cases, this number is unknown and certainly far larger than any of the previous categories.

For most experts, the conclusion is that there are adequate mineral resources for the foreseeable future. Falling production numbers should not be regarded as indicating resource depletion. There are two important caveats to this conclusion, however. The first is that production capacity and technology may be inadequate if there is a sudden spike in demand, and that transitions between production involving different forms of a particular material may be difficult. One example might be if hydrocarbon production has to transition from traditional oil fields to shale oil or oil sands, or eventually from petroleum to oceanic clathrates (methane deposits in ice matrices on continental shelves). Similarly, some materials tend to appear in low concentrations in ores and are thus produced as coproducts of other mining activities; bismuth is generally produced as a coproduct of lead mining, for example, and indium is a coproduct of silver and lead mining. For such materials, availability might be limited by the economics that drive the main ore-processing production cycle. Moreover, availability of many ore bodies is being limited by social, cultural, or demographic trends that put areas off limits to mining activity either formally or informally. In addition, any transition to new reserves usually implies lower concentrations of materials in ore matrices and higher environmental costs in processing: It could well be sinks for residuals, rather than simple material availability, which limit access to new reserves. Obtaining petroleum from oil sands, for example, requires more energy and water, and generates more waste, than traditional oil field production, all else equal.

will be permanently falling from this point onward. But, as Figure 4.4 shows, there are plenty of hydrocarbon resources. Estimates for global methane resources in gas hydrate deposits alone vary widely, but an accepted figure is 20,000 trillion cubic meters; for comparison, right now the United States uses just over 600 billion cubic meters per year. With uranium resources even larger given current consumption and the potential for renewables as infrastructure capability is developed (the **smart grid**), supplies of energy in an absolute sense are clearly adequate. Whether transition costs and delays, technology development, and management of environmental and social impacts can keep pace with demand smoothly is another matter.

Energy Source		Energy (ZJ)
Coal	Est. Resources	201.54
	Proven Reserves	24.83
	2005 Consumption	0.02
Oil	Est. Resources	11.14
	Proven Reserves	6.99
	2005 Consumption	0.15
Natural Gas	Est. Resources	11.52
	Proven Reserves	6.28
	2005 Consumption	0.10
Oil Shale	Est. Resources	17.16
	Proven Reserves	
	2005 Consumption	0.00
Clathrates	Est. Resources	500
	Proven Reserves	
	2005 Consumption	
Uranium	Est. Resources	2,184.00
	Proven Reserves	1.80
	2005 Consumption	0.04

FIGURE 4.4 World Nonrenewable Energy Data. Taken as a whole, fossil fuel resources are more than adequate for the foreseeable future. An important caveat is that limitations on use will more likely arise from the robustness of sinks (e.g., global warming) than resource availability. Also, the fact that hydrocarbons are potentially available does not rule out severe supply and demand fluctuations as transitions are made between different regulatory and technology regimes, and different forms of fossil fuel. (A transition to oil and petroleum, shale natural gas, or to methane clathrates, for example, could take many decades). A zettajoule (ZJ) is 10 followed by 21 zeros.

(Data from: Brown, C. E. *World Energy Resources*. New York: Springer, 2002; Organization for Economic Co-operation and Development [OECD]; "Unconventional Uranium Resources." Source: *OECD Nuclear Energy* 2006, no. 8 (2007): 182–192. Smil, V. *Energy in Nature and Society*. Cambridge, MA: MIT Press, 2008; United States Geological Survey World Energy Assessment Team. "U.S. Geological Survey World Petroleum Assessment 2000." http://pubs.usgs.gov/dds/dds-060/sum1.html#TOP: accessed August 29, 2009.)

Does all this mean that the sustainable engineer need not worry about material selection from a systemic perspective? Absolutely not. It is true that the data and existing analyses tend to suggest that questions of permanent scarcity for most materials are generally not a concern. However, any design that has the potential to stimulate significant jumps in demand for a particular material over a short period must be carefully evaluated by the engineer. Transition costs, environmental impacts and overloading of local or regional sinks for residues, strategic considerations such as control of material markets by unfriendly or kleptocratic governments, and short-term technical and financial barriers could pose problems that are just as severe as long-term availability. Thus, for example, it is likely that unless robust technological solutions are found for the increasing carbon dioxide in the atmosphere, which contributes to global climate change pressures, fossil fuel use will be limited not because it isn't available, but because the main sink—the atmosphere—isn't able to manage the increase in concentration of residues arising from fossil fuel use (primarily carbon dioxide). Similarly, availability of some minerals is more affected by where they are located and which governments control them than any questions of absolute scarcity.

4.4 FROM SUSTAINABILITY TO SUSTAINABLE ENGINEERING

Engineers are basically problem solvers. Whether working for a private firm, for a government agency, or as an independent consultant, an engineer's primary responsibility is to produce a solution that works in the real world, with all of the constraints—economic, competitive, technology availability, infrastructure deficiencies, ideological and political—that entails. Because the public relies on many engineered products, processes, and structures, virtually all engineering occurs in highly regulated environments. Globalization, with its mix of differing cultures, technological infrastructures, and regulatory regimes, adds to the complexity. To enable solutions in such complicated spaces, engineers are highly quantitative both by inclination and by methodological choice. Unlike many involved in discussions of sustainability, the luxury of constructive ambiguity does not exist for engineers: at the end of the day, the product, be it bridge or toaster or chemical production facility, must be physically instantiated, must work, and must meet all the different requirements of a very complicated design space.

One of the implications of being a problem solver is that one needs to change as the way that society defines its problem changes—that is, in fact, where the interest in sustainable engineering comes from. Unlike some past experiences, however, sustainable engineering, and the underlying anthropogenic Earth from which much of the complexity and concern arises, is not just a change in problem definition. It is also a change in how engineering is

conceptualized. Traditionally, engineers have been relatively instrumental in that they were presented with problems arising from design, manufacture or construction, operation, and management of technological artifacts of all types and scales. Both professional ethics and legal systems held them responsible for failure to carry out these tasks professionally and competently. Increasingly, however, engineers and technologists are being tasked with understanding the broader social, economic, and environmental implications of their work, with one implication being that they have some responsibility for those dimensions. Thus, for example, an engineering firm building a road might find itself responding to questions about how its road might change future settlement patterns in nearby sensitive locations, and engineers and scientists working on biotechnology or nanomaterial projects might find themselves being quizzed on the ethics of the underlying technology systems. This shift is a subtle change, but it has potentially huge implications for the engineering profession, leading directly to the question of how to take a young and somewhat confused social myth, sustainability, and interpret it in such a way that it can usefully inform and become a part of engineering practice.

Sustainability is often considered in three domains: economic, environmental, and social (this is often referred to, especially in the corporate context, as the "triple bottom line"). Increasingly, the last category is being viewed as too broad, however, and it can therefore be further broken into two categories: the social and the cultural. Although the two areas are clearly similar, the former focuses more on social status, roles within society, and groups and institutions characteristic of a society, as well as the relations among them, while the latter focuses more on underlying structures such as knowledge, beliefs and worldviews, art, morals, law and regulation, and custom.

What this breakdown does, however we choose to structure it, is provide categories that the sustainable engineer can then begin to address. Over the past two decades, mechanisms and supporting institutions have been developed that encourage progress towards sustainable engineering in all of these domains. Engineers have long had to struggle with economic factors; accordingly, engineering has developed many tools that help professionals include economic factors in their activities and designs. More obviously, corporate and project budgets and business plans exercise substantial power over any engineering project, and are therefore always part of engineering. Recent changes of focus reflecting sustainability considerations in engineering economics might include more emphasis on lifecycle costing methodologies, which enable initial design to reflect energy or environmental efficiencies that accrue over the lifecycle of the artifact or project. In terms of the environmental domain, many have suggested that sustainable engineering could benefit from the sets of tools that the field of industrial ecology has helped

generate, such as Design for Environment or life cycle assessment techniques (these tools will be covered in greater detail in later chapters). An engineer would thus first use the methodologies of industrial ecology, such as life cycle assessment, material flow accounting, or product and process matrix analysis, to determine environmental considerations, and then use existing design and engineering methods to integrate that knowledge into process, product, and infrastructure development.

Tools and methods to assess social and cultural domains are less well developed, in part reflecting the fact that issues in these areas are far more difficult to quantify and tend to be significantly more polarizing and subjective than in the environmental or economic domains. Industrial ecology professionals and "green engineers" have tried to extend their remit beyond environment into the social and cultural aspects of their work, but that is still a partial and somewhat controversial effort. Thus, what appears at first blush to be an easy solution—implement industrial ecology and achieve sustainable engineering—may be adequate conceptually, but oversimplifies the challenge to both industrial ecology and sustainable engineering in at least two major ways. First, industrial ecology and related fields almost always have arisen from environmental concerns so, not surprisingly, they may be relatively mature in enabling environmental issues to be identified, but they remain fairly primitive in addressing social and cultural issues. This not only reflects institutional evolution, but also the underlying truth that environmental issues are difficult to define and quantify; they can, however, generally be understood with some objectivity. Second, social and cultural issues are more subjective, contentious, and inchoate, and are therefore much more difficult to integrate into the quantitative tools upon which engineering and industrial ecology rely. For example, an engineer can with some confidence understand the environmental impact of creating demand for a trace metal that in turn engenders additional mining activity, but whether that activity can be considered "good" or "bad" from a social or cultural perspective is much less easy to determine. Some NGOs routinely oppose any mining; local communities, which may benefit from employment and local investment, often support responsible mining; the governments involved by definition support mining, or they wouldn't permit it (although in many areas different entities within the government may disagree strongly over who gets the spoils). Similarly, rules of thumb (heuristics)—such as designing for better energy efficiency during use, reducing toxic materials, and establishing a used cell phone takeback system—may enable one to design an environmentally preferable cell phone, but there are no heuristics for the social and cultural implications of a society that broadly adopts cell phone technology. Building Internet routers provides high-tech jobs and investment, but what the integrated effect of the Internet and its platformed services are on social institutions and cultures around the world is, at best, unclear.

By and large, then, many of the tools that the sustainable engineer will find available to this point will be either implicitly or explicitly in the economic and environmental domains, not the social and cultural domains. But this does not necessarily mean that one cannot say anything useful about the social and cultural aspects of a particular product, process, or project. First, the very fact that an activity is not only legal, but viewed by a government or a firm as worth undertaking, shows that there is some positive social or cultural value associated with it, at least from somebody's perspective. This point is frequently overlooked, in part because social activism tends to dominate discussions of social value, but it is nonetheless important to bear in mind. When constructing a new highway, for example, there may be strident opposition, but it is not immaterial that the political process has chosen to approve and fund the process.

Moreover, dialogs and disagreements about social value often reflect the fact that most engineering activities lead to differential distributions of costs and benefits: The groups that benefit from the project (say, truck drivers using a new highway) may not be the same groups that bear major costs (say, the people who are displaced so the highway can be built). This makes social values very difficult to quantify and almost impossible to smoothly integrate.

It does suggest, however, that a sustainable engineer should follow certain basic steps in considering the social and cultural dimensions of a particular activity.

1. Identify to the extent possible the social and cultural issues and values that the activity will affect, both positively and negatively. This is not as easy as it sounds, in part because engineers may not be good at empathizing with the various groups and communities affected, and in part because agendas may be hidden as well as explicit.
2. To the extent possible, create spaces within which dialog regarding various options can occur, so that social values that may otherwise not be identified or respected can be integrated into the decision-making processes. This can also enable identification of groups seriously interested in constructively working with the project and those that, for ideological or other reasons, will never do so and are simply trying to stop it by any means possible.
3. Identify changes in the project or design that can reduce identified social and cultural costs while at the same time meeting other relevant constraints (cost, consumer demand, and so forth).
4. Iterate these changes with the affected groups as a deliberate part of the project management or design activity.
5. Close on a final decision, and carry it out.

4.5 SUSTAINABILITY AND GLOBAL STATES

Sustainable engineers in particular need to be aware that sustainability in any particular application is contextual: that is, what may constitute an economically, environmentally, socially, and culturally appropriate design or project is not generally so much a function of universal principles, but primarily of local conditions. If a product is being created for developed-country markets, for example, it may be entirely appropriate to assume the existence of advanced infrastructure, such as potable water availability in the home, or sewage; if a that product is being designed for developing countries, those same assumptions might not be valid. For example, if a sustainable engineer is designing a manufacturing line for widgets and needs to include a chlorinated solvent cleaning system, that might work for a developed-country manufacturing environment, where employee safety standards and equipment will limit exposure to the chlorinated solvent (such materials tend to be neurotoxins and otherwise unhealthy), and regulatory and technological systems exist to ensure proper management of spent solvent when it needs to be handled as waste. In many poorer countries, however, employee safety standards are often lax in practice if not in law, and management of spent industrial solvents all too often involves dumping them in a local river. Under such conditions, the sustainable engineer should change either the process or product design to eliminate the use of the toxic solvent. (Why "toxic" is not necessarily "bad" is discussed in Box 4.3).

Values may also differ in different environments. Usually, it is not so much a question of different cultures having "different values," as much as prioritizing values differently. Thus, for example, surveys indicate that many people in developing countries are far more concerned with clean water and sewage than they are with global climate change, which tends to be the dominant environmental concern in developed countries. The reasons are obvious. It is not that individuals in developing countries are not concerned about climate change, but their health, and that of their families, is directly impacted in immediate and serious ways by improperly managed sewage and lack of safe drinking water. They place a much higher priority on such environmental issues than they do on climate change. More subtly, many people in developing countries place a higher value on consumption than activists would like, reflecting the fact that by and large they have less (often far less) than those in developed countries. Many Indians, for example, are eager to purchase the inexpensive small car, the Tata Nano, developed for Indian economic and infrastructure conditions ("the peoples' car from Tata Motors!"). Indian consumers view it as an opportunity to obtain for the first time what many Westerners have long had, personal automotive transportation (and they note that the Nano is safer than many of the current alternatives, such as wildly overcrowded scooters in heavy traffic). Western environmentalists, on the

BOX 4.3 Toxicity, Hazard, and Risk

Much of sustainability revolves around the idea of environmental quality and the safety of people involved in manufacturing products, using them, and disposing of them. It is therefore useful for the engineer, regardless of particular specialty, to understand the differences among toxicity, hazard, and risk. This is especially true because if this is not understood, fairly simple mistakes could be made, such as substituting a less toxic substance for a more toxic substance in design, but with the systemic effect of actually increasing potential environmental or human risk. Moreover, some NGOs will deliberately conflate toxicity with risk, even if there is no hazard, as part of their marketing campaigns.

Toxicity may be thought of as the inherent ability of a material to cause harm to a target system: an ecosystem (**ecotoxicity**), perhaps a human, or even a particular part of the human system such as blood ("hemotoxicity") or the nervous system ("neurotoxicity"). Different materials may be combined. To take one example, the Mojave Rattlesnake, which is indigenous to the desert where the author of this text lives, is the only rattlesnake to combine a hemotoxin with a neurotoxin.[3] For virtually all materials, there is a level under which harm is not caused, and a level over which virtually anything may cause harm. Water, for instance, will kill you if you drink enough too rapidly, and most venoms have levels under which they will not cause harm. A very few materials are will cause harm even with a single exposure—even a single asbestos fiber in the lungs, for example, creates a (very small) risk of asbestosis. Toxicity defined as the ability to cause harm is dose dependent: The more of a toxin you are exposed to, the greater the deleterious effect.

"Hazard" is not the same as toxicity; it requires that there be either a possibility or a pathway for a toxic material to interact with a target system. For example, a Mojave Rattlesnake in a completely unbreakable display case is just as toxic as one underneath your sneaker, but (assuming the display case really is unbreakable) it is not hazardous. Similarly, assume that you create a design using a toxic material, but where exposure to any target system is highly unlikely. You are using a toxic material, but the material in the design as you have structured it is not hazardous. This point is frequently not clear in regulatory regimes. For example, many countries heavily regulate leaded glass in electronics and other uses. Lead is a toxic material, but unless the glass is ground up and soaked in acid, it tends not to migrate from the glass within which it is embedded, reducing its hazard. This differentiation is important for many reasons. For example, if leaded glass is used to reduce exposure to radiation when a product is used and it is banned because it is "toxic," protection of the target system has been compromised, not assisted (at least until alternative technologies for radiation protection, assuming they are available, are implemented). On the other hand, leaded crystal is sometimes used for drinking glasses,

[3]To make matters worse, different subpopulations of the Mojave may lack the neurotoxin. The main lesson to be drawn from this is that, if you must get bitten by a rattlesnake, try not to get bitten by a Mojave.

decanters, and drinking mugs, and there is evidence that acidic materials, such as wine or coffee, can leach the lead from such formulations. In this case, there is both toxicity and hazard, and selection of alternative materials is clearly appropriate.

"Risk" is the probability of an event, usually but not necessarily considered negative, occurring. For example, if that Mojave Rattlesnake is loose on my hiking trail, it is hazardous; if it is sleepy, not sliding out of my way, and on the trail where I'm going to step, risk is high; but if it's off the trail and happily under a pile of rocks, it is still hazardous, but risk is low. Conceptually, this provides the broadest (scientific and engineering oriented) definition of risk:

$$\text{Risk} = (\text{probability of an event}) \times (\text{losses associated with the event should it occur})$$

Most engineers will get significant exposure to risk analysis and ways in which risk can be minimized, as part of their training; most of these treatments will be highly quantitative. Similarly, environmental scientists and toxicologists tend to use quantitative assessments of risk. It is important to understand that many people evaluate risk differently, however, and have a much more subjective and morally coded idea of risk. For instance, the quantitative risk posed by discarded chlorinated solvent in drinking water at low concentrations may be very small, but if the public believes that a firm behaved inappropriately, or if people think a cover-up occurred, they may well view the risk as significant and react very strongly, even if a toxicologist wouldn't. This subjective approach is not a "wrong" way to characterize risk, but it is different, and technologists especially need to be sensitive to the conditions that can encourage different perspectives on risk (such as morally unacceptable behavior by a polluter or a sensitive population at risk, such as children).

other hand, have generally attacked the Nano, because to them it represents an unacceptable contribution to global climate change.

More generally, as the first chapters of this book indicate, much of the challenge of sustainable engineering arises because the assumptions that underlie the way sustainability is framed—an egalitarian society, common closure on environmental values as dominant over other value systems—may hold for those in the sustainability discourse, but do not characterize the world generally. They are aspirational, not descriptive. Thus, it is possible—even probable given current trends—that, rather than heading for an egalitarian and high-quality environmental sustainability, we are moving toward a different form of sustainability. This different form would be characterized by increasing inequality between those with wealth and access to technology, and all others, where environmental quality is a designed function of developed societies and increasingly problematic in poorer societies. Figure 4.5 presents a list of some questions that are not relevant if we make the assumptions underlying

1. How many people at what average level of material well-being?
2. How many people with what distribution of material well-being?
3. How many people with what technology—and how rapidly is that technology changing?
4. How many people with what domestic and international political institutions?
5. How many people with what domestic and international economic arrangements?
6. How many people at what domestic and international demographic structures?
7. How many people in what physical, chemical, and biological environments?
8. How many people with what level of variability, stability, and resilience?
9. How many people with what risk?
10. How many people for how long? 150 years? Back to 50-year life spans?
11. How many people with what values, tastes, and fashions?

FIGURE 4.5 Which Sustainable World? This figure presents a partial list of variables that would affect what a "sustainable world" might look like. That many people would answer these questions quite differently depending on their cultural or psychological background is an indication of the danger of assuming that any particular formulation of sustainability is adequate for a complex and multicultural world.

many sustainability dialogs, but that, depending on the project or design the sustainable engineer is involved in, may become important for solving problems in the real world.

We end this chapter with several important observations. First, sustainability is a far more complicated issue for engineers than we might think on first blush. It has the characteristics of a social myth, but that does not mean it is not an important part of engineering; rather, it means it is critical that engineers understand and respond to sustainability's rather inchoate demands, now and in the future. Second, this movement towards sustainable engineering may, and should, draw on some existing frameworks, such as that of industrial ecology, but it will also require a lot more. It will require understanding more about the way that people think about both sustainability and sustainable engineering, and how engineers, working in the real world under existing constraints, can become the type of professional that their society requires and demands.

Exercises

4.1 You work for a large automobile company and have just come up with a breakthrough design for an internal combustion engine built from a ceramic that uses several kilograms of indium per engine. How would you evaluate the implications of this material use from the viewpoint of sustainable engineering? What kinds of questions would you want to ask as part of such an evaluation?

4.2 It is said that Alberta, Canada, has more oil in its oil sands than does Saudi Arabia, and many estimate that the methane resources available

in continental shelf clathrate formations are more than all other hydrocarbon resources—oil, coal, and natural gas—put together. Assuming these are reasonable first estimates of resource availability, why is anyone worried about fossil fuel use?

4.3 The World Commission on Environment and Development used equity as a platform for developing its (by now) widely accepted definitions of sustainable development. The Commission considered equity both between generations and within each generation. How does this matter from a practical standpoint?

4.4 You are the head of a design team that has been asked by your marketing department to create a "sustainable cell phone." Write a memorandum to your boss explaining how you intend to do that.

4.5 Should flying in a commercial passenger aircraft be considered a sustainable form of travel? Why or why not? Would it make a difference whether one is flying short-haul (under 1500 kilometers) or long-haul (greater than 3000 kilometers)?

Annotated Bibliography

1. There are a lot of books on sustainability and sustainable development. Most of them tend to tell you *what* to think, instead of *how* to think, and are thus of limited value. If you agree with them, they'll just reinforce your predilections, and if you disagree, they're not well-enough thought out or argued to enable you to learn from them. Among the ones that are worth reading is, of course, the "founding volume": *Our Common Future* by the World Commission on Environment and Development, also known as the Brundtland Commission, after the former prime minister of Norway, who led the effort. Much of what follows the 1987 release of this report, like much of what followed Marx's *Das Kapital*, pretty much fails to come up to its standards.
2. Sustainability and sustainable engineering have not yet built up a significant body of work, so there's not much to recommend along those lines. *Reconstructing Earth* by Brad Allenby (Washington, DC: Island Press, 2005) has some background on earth systems engineering and management that is relevant, and the newest iteration of the classic industrial ecology text by Thomas Graedel and Brad Allenby, *Industrial Ecology and Sustainable Engineering* (Upper Saddle River, NJ: Prentice Hall, 2010), has additional background, although the latter does not entirely escape from an environmental bias in its conceptualization of sustainable engineering. The 1998 book *Industrial Ecology: Policy Framework and Implementation* (by Brad Allenby; Upper Saddle River, NJ: Prentice Hall) covers the intersection between industrial ecology and policy, with discussions of various aspects of sustainability as a framework for much of the discussion. Note that all three of these books are authored by the same guy that wrote this book, and that

authors tend to agree with themselves, so take this recommendation with the appropriate caution.

3. Questions about mineral resources are varied and complex, and as one might imagine the literature is enormous. One of the best general sources is the *Minerals Yearbook* published by the United States Geological Survey (USGS). (There's usually a lag time of several years because obtaining and compiling the data takes a significant effort, especially given confidentiality concerns on the part of private firms.) It not only covers domestic production and consumption data, but it has articles on individual materials presenting national and international data. In addition, other national geological agencies, such as that of the United Kingdom, usually publish equivalent volumes. For energy information, perhaps the best thing to do is start by accessing the Energy Information Agency, or EIA, at www.eia.doe.gov. As with the USGS information, both national and international statistics can be accessed through this site.

CHAPTER

5 Homo Faber: Human History and Technology

A famine in petrol appears to be inevitable in the near future, owing to the fact that demand is increasing at a rate much greater than the rate of supply. . . . It is evident from the whole tone of the report that the committee expects to find in denatured vegetable spirits [alcohol] the fuel of the future.

—FROM *NATURE*, AUGUST 8, 1907[1]

5.1 INTRODUCTION

One way to write a chapter on human history and technology is to present a litany of different technologies and the dates and locations they were developed. This would be easiest for the author, but sadly uninformative for two reasons. The first and most important is that technology is far more than artifacts plucked from their social, cultural, and economic context. Second, such an approach is undermined by the very technology it purports to chronicle. Anyone with a computer now has access to search engines that can provide such a laundry list easily and in far more detail than the most comprehensive printed source. Such a chapter, then, would appear erudite but would in fact be superficial, even trivial.

[1]From *Nature,* August 8, 1907, quoted in "100 Years Ago," *Nature* 448: 653, August 9, 2007.

The real trick is to understand technology as a defining dimension of the Anthropocene, a part of what makes us human, perhaps what is even inherent in being human. Some argue, for example, that the strength of the human brain lies in its ability to integrate smoothly with whatever technologies extend the capability of the human animal, be it stone ax, a caravel sailing ship, or the hottest search engine. Technology taken in this sense is far more interesting and meaningful than simple lists of things and when they were invented. This approach, however, does highlight some basic issues. Most obviously, it forces us to grapple seriously with how we might choose to define *technology*, a term that everyone is familiar with and senses the meaning of, but finds difficult to define with any rigor. (Box 5.1 presents a working definition of technology for our purposes.) Nonetheless, a few fundamental points can be made. First, technology is more than simply artifacts, machines, and physical products; it is also the information systems necessary to operate them and the institutional framework within which they exist. A navigator on a ship, for example, may use Global Positioning System (GPS) technology in the form of a screen built into the ship's controls, but without training and an institutional infrastructure that supports the space program and satellite system that form the GPS infrastructure, there is no functioning technology. Behind that explicit GPS technology is a worldview that assumes that such measurements are valid and do not change arbitrarily over time, and that predictably moving from place to place in vehicles or ships is something that is important to do. Artifact, changes in individual cognition (in the form of training, for example), and institutional structure all must be present and work together to establish GPS as a working technology.

BOX 5.1 Aspects of Technology

Many people think of technology as hardware—an artifact, a product, or if the area is biotech, perhaps a designed DNA fragment or protein, or a biotech seed. But technology has always been more than that: Even in the Stone Age, to chip an ax head meant learning, skill development, and a cultural pattern that had a role for individuals with axes, and, some argue, the basic social structure that went along with the creation of and need for a stone ax. Accordingly, if we wanted a more inclusive definition, we might argue that technology is, quite simply, the means by which humans get things done, usually, but not exclusively, involving designed physical objects.

Such a definition implies a fairly broad view of what a technology might be, along the lines of Figure B5.1.1. It also hints at one reason why society is demanding sustainable engineers: Because technology is the interface between humans and natural systems, and increasingly involved in explicitly defining

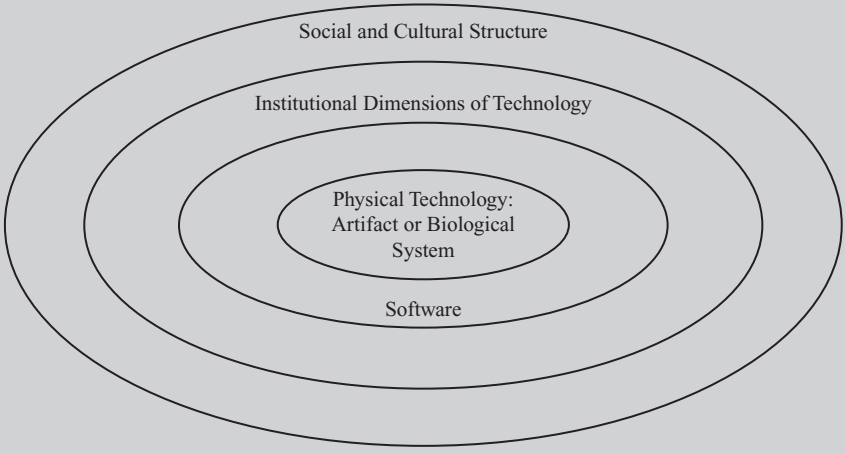


FIGURE B5.1.1 Technology System Levels. This figure depicts a schematic perspective on different levels of technology systems.

communities, and institutional and political structures (think of social networking sites), it is of growing interest to more and more people that it be designed, built, and managed responsibly.

Consider, for example, how one might think about a complicated technology system, such as a virtual workplace. At the most basic level would be enabling technology platforms: computers, wired or wireless routers, Internet connectivity, energy supplies, themselves built of complicated subassemblies and components. At the next level would be the information structures that knit together the various artifacts. (Whether that includes the software with the hardware level is a relatively arbitrary choice; I have separated them because software can raise an entirely new and complicated set of issues, such as security against hacking.) At the next level would be the institutional designs necessary to support virtual workplaces: this might include, for example, making sure that the company's data and critical information were online, secure, and accessible to those who needed it to perform their function. Managers would have to be trained to supervise virtual employees; lawyers would have to make sure relevant laws, most designed for office building environments, were followed. Finally, at the highest level one would find the affects of virtual workplaces on social and cultural practices, and visa versa. Thus, for example, it might be that cultures that highly value face-to-face contact, such as in Japan, are less open to virtualizing the workplace than elsewhere. Alternatively, virtual workplaces might be more accessible to physically challenged, rural, or older workers, and thus lend themselves to greater equity in economic opportunity.

This example illustrates several principles for the sustainable engineer. First, it is important to understand the level of the system at which we are working, and how the constraints and opportunities present are affected by other levels. Designing security for a corporate network, for example, would be more challenging if a firm

were virtual, and thus its network were more "open," than if computing assets were all under the physical control of centralized corporate facilities. Second, no physical or software technology exists in a social or cultural vacuum; although in many cases social input is standardized and relatively routine (e.g., safety standards for consumer designs), in other cases it increasingly is not, which in turn requires engineers who are more able to respond to unanticipated stakeholder demands and actions. Finally, engineering in today's world can no longer adequately be done by professionals who view themselves as strictly focused on the artifacts and hardware of technology. The complexity of these systems, and increasing interactions between the physical and social dimensions of technology, require that engineers prepare themselves to be policy and business leaders, rather than simply well-trained technicians.

A few moments' thought will confirm that, for the sustainable engineer, this is a comfortable definition. Few designs do not incorporate cultural dimensions, even if they are as prosaic as safety requirements for electrical appliances or sharp tools, or aesthetic and ergonomic considerations for consumer goods. And, of course, the institutional aspect becomes relevant every time there is a budget limit for a particular project or design. Beyond these concerns, however, technologies are potent means by which humans perceive, define, and manipulate their environments, and are thus major contributors to the world as we know it. (An obvious case in point is how the internal combustion engine and automobiles have called forth not just transportation networks and specific patterns of built urban infrastructure, but also determined important dimensions of atmospheric chemistry and physics, as well as creating automobile-oriented cultural patterns.)

In some very important ways, the history of humanity is coupled to, and coevolves with, the history of technology. Indeed, as Figure 5.1 illustrates, the three major eras of premodernity, the Stone Age, the Bronze Age, and the Iron Age, are named for the dominant technologies by which they were characterized. This labeling also nicely illustrates the point that, until fairly recently, dominant material, dominant technology state, and dominant cultural patterns were fairly similar at a high level. (This observation should not be taken to be **technological determinism**; see Box 5.2. It can also be understood as another way to frame the relationship between culture, population level, and technology state, as suggested in Figure 1.5.) This relationship, in fact, lasts until what some scholars call the "Technolithic," when intellectual capital and software comes to dominate the material foundation of culture (and when material use becomes far more complex, as many more materials are integrated into economic activity).

Age	Dates (approx.)[1]	Social Context	Technology Context	Dominant Material
Eolithic (Dawn Stone Age)	Ca. 10M ybp	Evolving bands	Origins of toolmaking	Wood, stone, and bone
Paleolithic (Old Stone Age)	5M to 12K ybp	Hunter–gatherer	Hand axes and shaped materials	Stone
Mesolithic (Mid Stone Age)	12K to 7K ybp	Hunter–gatherer	Origins of blade technology	Stone, clays
Neolithic (New Stone Age)	6K to 3K ybp	Agrarian revolution/towns	Beginning of copper use	Stone to metal
Chalcolithic (Bronze Age)	3K to 1.5K ybp	Early urbanization based on agriculture	Tin (therefore bronze) discovered ca 3000 ybp	Bronze
Iron Age	1.5K to 500 ybp	Development of urban civilizations	Development of technocultural systems, *homo faber*	Iron
Technolithic	500 ybp to now	Urbanization of species and posturban (virtual) societies	Industrial technology leading to NBRIC (nano/bio/robotics/ICT/cogsci)	Specialty materials, nanotech culminates control of material world

[1]In this table, M signifies one million; K signifies one thousand; ybp is an abbreviation for "years before present."

FIGURE 5.1 Human Technological Evolution. Until recently, technological evolution (in historical terms) was definable by the dominant material characterizing the period: stone, bronze (tin and copper), and iron. In fact, for almost all of human history stone has been the dominant material and technology base underlying human society. That role began to end as the agricultural revolution and consequent urbanization began to characterize human cultures. Only with the flowering of the Industrial Revolution, however, has intellectual capital rather than dominant material come to define technology state, a process that has shifted into high gear with the Five Horsemen of emerging technology: nanotechnology, biotechnology, robotics, information and communication technology, and applied cognitive science (nano/bio/robotics/ICT/cogsci, or NBRIC).

(Based on: McNeil, Ian, ed. *An Encyclopedia of the History of Technology*. London: Routledge, 1990; Freeman, C. and F. Louçã. *As Time Goes by: From the Industrial Revolutions to the Information Revolution*. Oxford: Oxford University Press, 2001.)

BOX 5.2 Technological Determinism

There is an important difference between understanding the power of a particular discipline or domain of human activity, and concluding that, taken alone, that single approach explains human history, or is able to predict the future. It is very important that sustainable engineers understand the very potent effects of technological systems, which are almost always underestimated by policy makers, scientists, environmentalists, and social scientists; but it is equally important not to fall into the trap of thinking that technology "determines" either the past or the future. It is part of the complex evolving system that is our anthropogenic Earth, but it does not define or determine that system. Thinking that it does is an intellectual mistake called "technological determinism"—that is, it is the mistake of thinking that technology alone determines the shape of cultural, social, economic, religious, and all other human systems.

Technological determinism is by no means the only or even most familiar form of determinism. It is not uncommon for religions to claim that their deity is omniscient and omnipresent. Perhaps the most famous secular determinism is the economic determinism of Marx, who essentially held that the "means of production" determined cultural, institutional, and social forms. Thus, for example, industrialization created a class of capitalists, and a larger class of proletarians, and the conflict between these classes, which arose from economics, then drove history. It is worth noting that Marx derived his economic determinism from what he considered an Enlightenment, scientific approach to history, and that it is quite possible that he would have modified this view in light of subsequent history. Others have more recently been called advocates of "environmental determinism," in that they argue that environment is dominant in determining social, cultural, institutional, and economic form. Some reviews of Jared Diamond's book *Collapse*, for example, felt that his strong emphasis on environmental factors as determining the fate of societies slipped from rigorous analysis into environmental determinism. Although this seems to overstate Diamond's position, it is clear that the more popular environmental issues become, the more environmental determinism becomes a socially acceptable (if completely inadequate) way of understanding a very complex world. The idea that carbon footprints are adequate to capture the costs and benefits of products, for example, is clearly wrong, but can be seen as a particular form of environmental determinism (the single-minded focus on global climate change and carbon footprints has been called **carbon fundamentalism**). Some scholars of military activity seem to verge on **military determinism**, which sees war, and the militarization of society, as determining human history as a whole.

All determinisms work because they capture a part of a very complicated reality. Using the understanding of complexity developed in Chapter 3, we could say that each perspective can call forth partial, but useful, perspectives from the complex underlying reality. Religious determinisms play on the very real, and strong, psychology of human faith and its role in social and cultural conflict; economic determinism emphasizes the critical and undeniable role of economics in

structuring societies; environmental determinism emphasizes the importance of the interactions of the human and built environment with the natural environment in a period when such regional and global linkages are becoming more apparent and more important (think of climate change); military activities frequently mark turning points in human history (and there is probably little question that an all-out nuclear war between the Soviet Union and the United States would have been pretty determinative of a lot of things). But all determinisms are partial. The mistake, therefore, is not in highlighting the power of a particular domain of human endeavor or thought, but in extending that domain beyond the boundaries within which it is valid. Religion, economic class, military activity, and yes, technology, all may exercise critical power at certain points in our history, but they are all in fact coupled. What we view as an "economic" development is, in fact, a multidimensional phenomenon, but we are choosing to highlight the economic aspects of the phenomenon.

Technology, then, is a powerful, underappreciated, and integral part of being human and of human history on this planet. It does not, however, "explain" or predict human past, present, or future, except in a partial and contingent sense. And, as the railroad example in the text clearly illustrates, the more powerful a technology, the more it coevolves with other domains: institutional, economic, social, cultural, environmental, and the like. Kondratiev waves are useful devices to pull technology systems out as a critical element of a complex underlying system, but they are a way of perceiving history, not predicting it.

Moreover, that technology is such a strong dimension of human activity, and so obviously affects the Earth and its biosphere, does not mean that it can solve all problems. The idea that technology will provide fixes to our complicated problems—which could be called **technological optimism** rather than technological determinism—simply fails. Technology alone cannot make cities more livable, raise the global poor from poverty, win conflicts in Afghanistan or the Congo, or feed the hungry and provide fresh water to those currently lacking it. Technology may well be necessary to successfully meet such goals, but it is not in itself sufficient. We work with technology because we are engineers, but we are sustainable engineers to the extent that we can design and engineer within the broader and more complex context within which modern technologies evolve—not underestimating the power of technology, as many do, but not overestimating it, either. Technology is a powerful force, but it does not determine our future.

The history of many technology systems in some ways mirrors this pattern as well. For example, the history of money as a technology, summarized in Box 5.3, also displays the evolution from an emphasis on a few natural materials (stone or shell, then easily found elemental metals such as gold and silver) to a high-technology state (e.g., dematerialized credit card transactions). Energy evolved from wood, to coal, to processed fossil fuels, to secondary

carriers such as electricity characterized by increasing fragmentation of production technologies with the resultant power distributed over what is becoming a smart grid. The substitution of intellectual and built capital for labor over time even extends to military technologies, especially in the last century—a trend becoming quite apparent with increasing reliance on unmanned robots in combat zones, and migration of conflict to cyberspace. Thus, we can identify three important trends in technological evolution:

1. Evolution towards greater complexity
2. Evolution towards increasing systems integration, and thus increasing system scale and scope
3. Continuing substitution of capital, and increasingly of information, for physical labor

But this only begins to scratch the surface of how technology evolves and affects us, a subject perhaps best presented through consideration of a familiar technology, railroads.

BOX 5.3 The Technological Trajectory of Money

Money is both an important technology in itself, and a necessary, if not sufficient, enabler for technological evolution. Without money, research, innovation, and deployment of new technologies are difficult if not impossible. New technologies and technological infrastructures, such as the "smart grid," are very expensive to design, build, and maintain. Financial constraints are a part of virtually any engineering activity.

But money is also an interesting case study in technology and how it evolves over many centuries. Figure B5.3.1 shows a schematic of the evolution of money, from its early roots as a decorative and thus valuable physical artifact to its current instantiation as, well, insubstantial.

The "technology of money" illustrates several important trends. First, money evolved from a value primarily based on its physical attributes (shells; iron in ancient China; gold and silver), to one based primarily on social and cultural attributes. Modern paper money in a country such as the United States, for example, is based on the fact that it is "backed" by the federal government or, more broadly, on a widely accepted faith that the government will not default on its money. A dollar bill is essentially a contract that we take on faith; it is a legal construct, with no inherent physical value. (Try to sell the paper in a dollar bill and see how far you get!) Even more now, however, money has become virtualized: When you purchase something physical with your credit card, the relevant transaction occurs somewhere in cyberspace. If computers crashed, so would the worth of most Americans, because it is built on legal concepts, electrons in space, and faith in the underlying strength of the economy—which, in a nice bit of circular logic, depends on that faith in money to stay liquid.

Type of Currency	Scale of Validity	Economic Structure Supported	Ratio of Material to Information[1]
Decorative (shell)	Within local culture	Barter	High M
Coin (certified composition)	Regional	Regional trade networks	High M, but I important
Coin (certified value)	Regional	Regional trade networks	Increasing I; value from inherent to cultural (value of metal used doesn't have to equal economic value)
Paper currency	Global, depending on issuing entity and requiring interchangeability	Global trade networks; sophisticated economies supported	High I; M content becomes virtually irrelevant
Electronic	Global/cyberspace	Most complex	Money dematerialized; becomes information function

[1]In this column, M signifies Material; I signifies Information.

FIGURE B5.3.1 Evolution of Money. This figure shows the evolution of money from decorative, to physical, to representational, to virtual. In each guise, the "technology of money" changed profoundly, as did the technology (and cultural and economic) state that each form of money was able to support.

It is, of course, true that to some extent money has always been a cultural construct. Economists have argued for many years over exactly where the value of gold or silver actually comes from—why, in other words, should it be a physical instantiation of value, even in economies where markets in materials are unfamiliar? Some of that answer is that enough people began to value gold and silver so that these substances became accepted means of identifying, and eventually quantifying, value; in other words, money from the beginning has had a social dimension. In this, it is not unlike any technology.

In addition, as money has evolved, it has coupled with a number of other technologies, some tightly, and some not so much. For example, it couples tightly to the technologies required to make it. Many years ago that might have been large stamping presses in mints that made coins; today, it is much more a function of sophisticated software platformed on complex institutional and technological systems. Most people, for example, don't hold "money" even in electronic form: they hold stocks, bonds, derivatives, and other financial instruments that didn't exist (at least not in anything like their present form) 200 years ago—and those financial instruments are increasingly virtual as well.

More subtly, money couples to other institutions and technology systems as a necessary technology. Arms races and national competition, for instance, are nothing new, and they have always required money. It is no surprise, then, that financial institutional innovation (creation of a central bank and transparent government accounts enabling institutionalized low-cost loans, for example) gave the British a lead in funding over the French in the 1700s. The Bank of England was founded in 1694, and perhaps the most important person in European history of whom no one has ever heard is the Chancellor of the Exchequer from 1694 to 1699, Charles Montagu. Montagu oversaw the establishment of the Bank and the development of the beginning of modern state financial institutions. The Bank and the financial strength it created for the English contributed in no small part to the British ability to field the world's most potent blue water navy over that period, which in turn contributed to the country's globalized imperial empire, which in turn contributed more trade and more money—and to British victories over France over the next century, culminating in the defeat of Napoleon in 1815. A more modern example is venture capital, an important part of the innovative process, and an area in which the United States, with its acceptance of free market principles and libertarian bent, excels.

Finally, the changes in the technology of money do not just make money itself more complex (e.g., derivatives that get so sophisticated no one can keep track of financial risk anymore, including the firms that made them). Rather, an accelerating evolution of money technology was necessary to support the rapid evolution of other technology and social systems. Thus, for example, if we were still dependent on gold and silver money, it is not just that the process of buying a car would be painfully physical (wheelbarrows full of silver dollars?); it's that the complicated supply network required to build the car in the first place couldn't operate. The large industrial combines of the age of industrial growth in the United Kingdom and United States could not have been funded through physical transfers of metal money. Today's globalized economic structure would certainly be impossible if money were still physical, as illustrated by the panic of 2008, which in effect drained much electronic credit from the world's systems and thus almost precipitated a major global depression.

In short, money is an essential technological infrastructure, and its evolution towards more complexity over time is both a story of technology in itself and absolutely necessary if other technology systems are to continue to evolve towards more complexity. So from a sustainable engineering perspective, we could do much worse than to begin by understanding money.

5.2 THE RAILROAD AS TECHNOLOGY SYSTEM

It is one thing to observe that technology systems should be understood as integral parts of the human/natural/built systems that characterize the anthropogenic Earth, not just as a collection of artifacts. It is another to understand

how powerful these technology systems are in their ability to shape people, firms, and society as a whole, as well as the complexity they introduce as they leave the laboratory and the pilot test phase, and begin interacting with human and social systems, which have a far different, and higher, degree of complexity. Remember also that, although we focus on technology in this discussion, all these systems are complex and thus coevolve; it is not that mass production of consumer goods such as cars and refrigerators "caused" mass credit and mass consumption social patterns, but that such things evolve together. The technology focus suits us best in the following discussion, but a financial historian might just as easily focus on the financial changes associated with major technological systems. These features are perhaps best illustrated through historical example.

Let's take railroads for an example. Today, railroad technology is a familiar, even boring, element of the human environment, eclipsed by air transport and networks of automobile and truck traffic. But in the middle 1800s when the technology first began to be rapidly adopted, the railroad was not just the most impressive piece of machinery most people ever saw: It was a sociocultural juggernaut. It changed the world completely, and in ways that no one foresaw, or perhaps even could foresee. The world before the railroad, and after, was different in profound ways. This was especially true in the United States, where by far the most rail was laid in this period, and where the long distances between settlements, and sparse settlement, made rail an especially radical change in transportation technology. Figure 5.2 illustrates both the rapid takeoff of this technology and how much of global rail construction took place in the United States. Early figures demonstrate how rapid the technology diffused in the United States. In 1830 there were only 73 miles of track; by 1840, 3,328 miles of track had been laid; by 1850 this had more than doubled to 8,879 miles of track; while between 1850 and 1860 the mileage jumped by almost a factor of five, to 30,636 miles. Figure 5.3, however, indicates that on a density basis—miles of track per unit area—the differences in railroad technology adoption between the United States and most European countries is almost inverse, with Belgium, a relatively small country, laying more track per unit area than any other nation.

For the sustainable engineer, perhaps the most salient effect of the railroad is the way it created the need for, and was dependent on, coevolution of coupled technologies. Conceptually, railroads are networks that operate in real time, so of course they required a uniform, precise system of time that was coextensive with the rail technology. Railroads were thus an important factor in the creation of "industrial time" and its associated culture. Before railroads, local times were isolated: London time, for example, was four minutes ahead of Reading, over seven minutes ahead of Cirencester, and fourteen minutes ahead of Bridgewater. In the United States around 1850, there were over 200 local times that differed from each other in relatively arbitrary ways. The reasons for such patterns were very simple: connections between these nodes (towns)

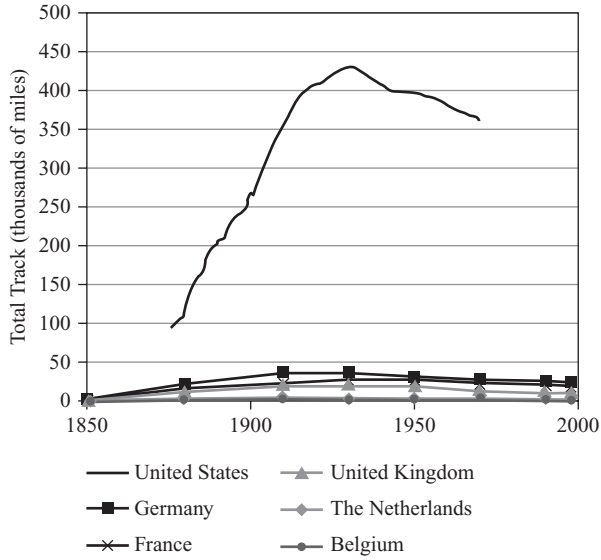


FIGURE 5.2 Railroad Diffusion. This graph dramatically illustrates the growth of the railroad in major industrialized countries of the time, but especially in the United States, where both the expansionist national culture ("Manifest Destiny") and economic needs drove rapid adoption of the technology.

(Data from: Fremdling, R. "European Railways 1825–2001: An Overview." http://ggdc.eldoc.ub.rug.nl/root/WorkPap/2002/200254: accessed September 3, 2009; United States Census Bureau, S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright, eds. *Historical Statistics of the United States on CD-ROM: Colonial Times to 1970*. New York: Cambridge University Press, 1997.)

involved walking, horses, or in some cases, canal boats, and thus time coordination was defined by the time cycle of such technologies. Ten-, twenty-, or thirty-minute slippages between nodes were simply immaterial, because the times involved in transit between the nodes were much longer.

With trains, of course, minutes mattered (or else you had trains bumping each other at high speed, a highly undesirable behavior), so time technology had to change in response. (Similarly, the networks on your computer chips are by far faster, so the time networks connecting the nodes on those chips need to be correspondingly more accurate.) But time, of course, is a heavily cultural technology, so the adaptation to uniform systems of time was neither smooth nor instantaneous. For example, for a considerable time in the United States, each train company had its own time, so that stations serving several train companies had different clocks (Buffalo had three different clocks at one point, Pittsburgh six). In fact, regional standard time did not gain legal recognition in the United States until 1918. Thus, another lesson from the railroad/time nexus: the cycle time of technological adaptation is moderated not just by research and engineering, but by social and cultural adoption of the necessary elements of technology, and the latter can take quite a while.

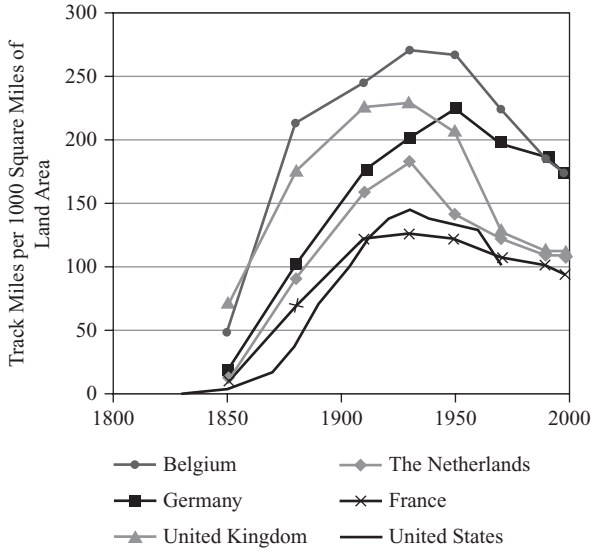


FIGURE 5.3 Railroad Density. This figure shows the diffusion of rail technology in terms of track miles per 1000 square miles. A comparison of Figures 5.2 and 5.3 shows that the absolute amount of track laid in the United States was far greater than elsewhere, but that was a result of the fact that the Unites States is far bigger geographically than the European countries.

(Data from: Fremdling, R. "European Railways 1825–2001: An Overview." http://ggdc.eldoc.ub.rug.nl/root/WorkPap/2002/200254: accessed September 3, 2009; United States Census Bureau, S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright, eds. *Historical Statistics of the United States on CD-ROM: Colonial Times to 1970*. New York: Cambridge University Press, 1997.)

But time is not the only critical network function for something like a railroad; there must also be some communication technology co-extensive with the network. Before the railroad, letters and other documents worked just fine: carried by horse, or person, or carriage, or sailing ship, the speed of a letter was exactly that of the dominant transportation technology: nothing faster was required. In the case of the railroad, however, coordination of railroad operations by letter carried by train would clearly not work; the more rapid communication system needed was provided by the development and deployment of telegraph technology. Not only was telegraph technology co-extensive with rail networks—frequently laid along the same rights of way—but it was an absolutely necessary coevolved technology infrastructure for the creation and operation of regional integrated rail systems. Again, however, the introduction of these technologies was not absolutely in sync; different parts of co-evolving systems usually evolve at different rates. This can be important if one necessary subsystem evolves at much slower rates than other parts, because that can act as a constraint on rapid innovation. We can certainly build plug-in vehicles, for example, but creating appropriate supporting energy infrastructure will take a much longer time.

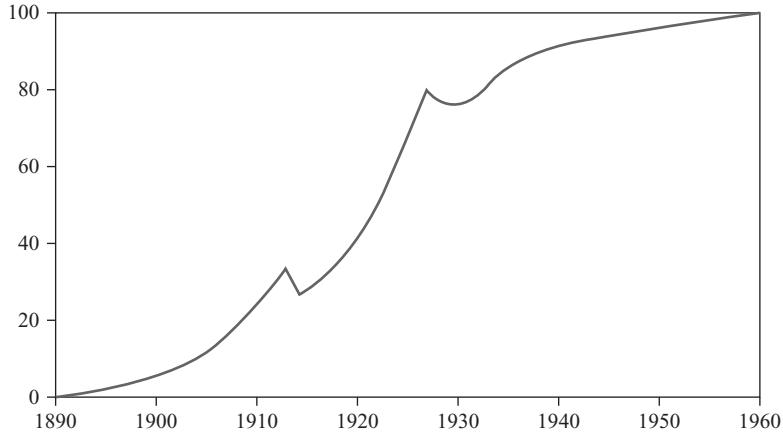


FIGURE 5.4 Crossties Treated with Chemical Preservatives, United States. This figure shows the rapid growth in the use of chemical preservatives in railroad crossties, a technology that alleviated unsustainable demand for timber on the part of the railroad industry, which was itself undergoing exponential growth over this period.

(Based on: Ausubel, J. H. "Regularities in Technological Development: An Environmental View," in ed. J. H. Ausubel and H. E. Sladovich, *Technology and Environment.* Washington, DC: National Academy Press, 1989, pp. 70–91. Reprinted with permission.)

Railroads also had substantial environmental impacts. Most obviously, they required coal and wood for fuel; tracks also consumed vast amounts of wood, to the point that Teddy Roosevelt argued in a famous speech to the American Forest Congress in 1905 that

> Unless the vast forests of the United States can be made ready to meet the vast demands which this [economic] growth will inevitably bring, commercial disaster, that means disaster to the whole country, is inevitable. The railroads must have ties . . . if the present rate of forest destruction is allowed to continue, with nothing to offset it, a timber famine in the future is inevitable.

As Figure 5.4 indicates, technology did improve—in this case, the development of wood preservatives, particularly creosote, that enabled far less use of wood. This example makes two further points of interest to the sustainable engineer. Most important, just because something is renewable doesn't mean it is necessarily better economically, socially, or environmently. Especially when scales are large, renewables can be just as problematic from an environmental perspective as nonrenewables, perhaps even more so. Second, although it is never a guarantee for the future, it is at least a robust argument that, for the most part, when environmental problems have been identified in the past, technology has provided solutions. Each solution may have generated new issues in its own turn, but solutions they were. An analogous situation to the

railroad ties, for example, was the English problem in the 1600s as usable timber for the navy and commercial shipping, which was the core of their growing power, began to run out. In this case, it was the colonization and exploitation of the New World, with its easily available stores of timber and pine tar, that helped resolve the potential scarcity problem—a solution that relied on advanced European ship and military weapon technologies (and in turn fed those technology systems).

Interestingly enough, the technology developed to preserve railroad ties, creosote, is not approved by regulators for home use because of the potential hazard. ("Wood creosote," made from high-temperature treatment of woody material, has a different composition than creosote derived from coal tar—the latter is far more toxic because it contains high concentrations of polyaromatic hydrocarbons and is considered carcinogenic (cancer causing). It is coal tar creosote that is used to treat rail crossties). Nonetheless, many used creosote-treated railroad ties end up being used in residential gardens because . . . surprise! . . . they don't rot. When they are retired from use by railroads, therefore, they are not disposed of, but instead are recycled into landscape uses, with management and exposure implications quite different than those arising from use in the railroad context. This example offers another lesson for sustainable engineering: That something is designed to be used as an industrial product does not mean that it, or a by-product, won't have a new life as a consumer product, being used in ways for which it was neither designed nor intended. (Something similar from a green chemistry perspective occurs when antibacterials pass through sewage treatment unchanged, and thus become a presence in surface water—in effect, the chemical engineers involved are designing not just a consumer product, but surface water chemistry.)

But this was by no means the only environmental impact of railroad technology. Railroads transformed landscapes at all scales. In the American Midwest, for example, Chicago arose as a major city because of railroads. Moreover, by becoming a major agglomeration and transshipment center for agricultural production, Chicago and the railroads for which it was a major node in turn changed the entire Midwest from swamp and wilderness into a gridded and farmed landscape. The dramatic and complete change in land use was enabled by the railroad, and would have been impossible without it. It is not too much to say that much of the North American landscape has been designed by the railroad and associated policy decisions (such as giving railroads land on each side of their routes to encourage them to lay track faster and further than they might have otherwise).

The changes induced by railroad technology were not, however, just environmental or technological. Railroad firms played a large role in shaping modern managerial capitalism. Railroad technology, because it was so much bigger and complex than factory technology, required an entirely different structure

for the railroad firm. Whereas previous industrial technologies had called forth a division of labor among "blue-collar" factory workers, the scale of railroad enterprises required a division of labor at the management level, with modern and professional accounting, planning, human resources, and administrative systems. Similarly, railroad firms were influential in shaping modern capital and financial markets. (Railroad construction was the single most important stimulus to industrial growth in Western Europe by the 1840s.) Because individual factories were the unit of production at the beginning of the Industrial Revolution, the capital required by the early factory system could be provided initially by aristocrats and landowners, and subsequently by factory owners using their own capital. This essentially individualistic financial structure was nowhere near adequate to support the huge capital requirements of railroad firms, and so new forms of capital aggregation—modern stocks and bonds—were required. Here also lies another lesson for the sustainable engineer: finance always matters. It may be just cost constraints on a particular project but, particularly where new technology systems are involved, it may be something a little more fundamental. So, for example, electric plug-in vehicles may need a smart grid, and a smart grid will require substantial financing, just like the railroad infrastructure did. The sustainable engineer working on plug-in technology should be sensitive enough to systems implications to be asking that question seriously.

As with environmental effects, the economic effects of railroad technology were also broader than just firm organization and finance. Like most major technological systems, railroads fundamentally changed national economic structure, and therefore political and social power bases. In particular, railroad infrastructure enabled the emergence of scale economies in core industrial sectors, and created viable national, rather than local, markets for commodities. When transit was by carriage, it was prohibitively expensive to ship lower value commodities over any great distance; even where canals were available, anything that might spoil couldn't be shipped for great distances because of the times involved. Railroads changed this completely. Before railroads, local businesses predominated in virtually all sectors of the economy, but the monopolies and trusts of turn of the century America—the sugar and tobacco trusts, Big Steel, Standard Oil, the railroads themselves—rode in on the new rails of steel. As railroad technology expanded, economic power also shifted to industrial firms from agriculture; more subtly, so did cultural authority. Thus, one effect of the railroad in the United States was to validate the continental scale of the American state and the doctrine of Manifest Destiny (**Manifest Destiny**, a term first popularized in the 1840s—not coincidentally, the decade when railroad technology took off in the United States—refers to the American political doctrine that the United States was intended by God to extend from the Atlantic to the Pacific oceans). But, of course, railroad

technology not only validated Manifest Destiny, but made it physically plausible: in a real and fundamental sense, railroad infrastructure was the physical network upon which the American state could be realized not just philosophically, but in fact.

It is also important to remember that most technology systems of any significance have military implications as well. Indeed, the Bronze Age is named in part because of the military dominance that bronze weapons provided over any previous military technology. So, too, with railroads. Much of the track laid in the early and middle 1800s in the United States, for example, was in the industrial North, and some historians have suggested that at least part of the credit for the outcome of the Civil War should go to the disparity in rail infrastructure and subsequent industrial efficiency between the North and the agrarian South. At roughly the same time in Europe, the Prussians and then the German state were among the first to recognize the importance of rail for military mobilization; rail stations near the French border in particular were designed with mile-long platforms so troop trains could unload rapidly, speeding the logistics of mobilization. Similarly, railroads helped reduce the importance of a major constraint previously hindering armies operating in the field: logistics. Napoleon's army, for example, foraged for its food and animal feed, as did everyone else at the time; that changed when railroad technology could ship materials as readily as it shipped the troops to use them. Railroad technology, in short, is "dual use" in that it is both a civilian and a military technology. In the same manner, highway networks such as the Interstate Highway System in the United States were built in part because of the increased efficiency of military operations they enabled, and modern ICT networks are deeply implicated in military operations and national security functions, both as sources of asymmetric military capability and as sources of vulnerability to cyberattack. Very few powerful technology systems of any kind are independent of military implications; it is always a useful exercise for the sustainable engineer to think a little about the potential military implications of whatever technology she or he is involved with, for they usually exist.

Every great technology system has psychological and cultural effects that are both profound and subtle, and the railroad was no different. Psychologically, railroad technology did not just extend, but obliterated, the sense of place and rhythm that previous transportation technologies, the horse and carriage and canal, had encouraged. Early reports of railroad travel frequently emphasized the psychological dislocation induced by the increase in speed and reduction in perceived distances that the technology introduced; the sense of velocity and time could not change so dramatically without having a significant impact on people used to a much slower pace. In this, railroad technology shows a somewhat uncanny similarity to the way people unfamiliar with ICT, social networking technologies, and multitasking are reacting to

technological change today. Another complaint frequently made about the railroad is that it converted people into commodities: in some ways, this is the psychological dimension of the first really efficient people-moving technology. Railroads enabled fast, quick, mass movement of people far more safely over far greater differences than any previous technology (compare the transnational railroad in the United States to the covered wagon that it superseded), and in doing so profoundly disturbed the way people had learned to think of themselves and their environment. Early rail travelers complained about feeling like mere packages in a delivery system, not unlike the way some air travelers today complain about low-cost airlines.

In some ways, for many people (especially in the United States, where the technology was so rapidly adopted, and which tended to frame both public debate generally and technology issues specifically in religious terms), the railroad carried religious meaning as well. For some, being able to create and manage the power and ferocity of a massive technology like the railroad represented a desirable human ascent towards salvation and God. To others, of course, the railroad was quite the opposite; as an Ohio school board fretted in 1828, when the technology was still new and to many people beyond the pale of what humans should presume: "If God had designed that His intelligent creatures should travel at the frightful speed of 15 miles an hour by steam, He would have foretold it through His holy prophets. It is a device of Satan to lead immortal souls down to Hell."[2] The railroads thus strengthened a coupling between religion and technology that, although very complex in the ways it plays out, continues in the United States into modern times. On the one hand, those that believe technology is taking humans into areas where religion indicates they should not go support policies such as stem cell research limitations. On the other hand, consider Robert Oppenheimer's famous musing at Trinity Test Site at White Sands, New Mexico, in 1945 at the first successful nuclear test, when he echoed Vishnu from the *Bhagavad Gita*: "Now I am become Death, destroyer of worlds." Notice in this last example the critical shift that technology enables: the aspect of destroyer of worlds in the *Gita* is divine; at White Sands, the divine has become human. With human technology may have come the power of the gods, but, as with nuclear weapons, that is a serious and potentially catastrophic power to vest in human institutions.

Culturally, the railroads changed the worldview that had heretofore defined Americans. Before railroads, the ideal American was the agriculturalist, bravely converting a wilderness landscape to farming (as many believed God intended). The ideal image of America that emerged from this worldview was a nation of independent, sturdy, yeomen working their own

[2]Ohio School Board, 1828, quoted in D. E. Nye, *American Technological Sublime* (Cambridge: MIT Press, 1994), p. 57.

land—Jeffersonian agrarianism. But with the railroad, Americans began to elevate a different worldview. The agrarian Edenic vision was increasingly augmented, if not entirely replaced, by one built on technology: factories, steel, oil, railroads—a high-technology road toward the New Jerusalem. Rather than simply managing farms, Americans began to see themselves as building their own world, a better world. This may sound relatively trivial, but it leads to very different policies; it is a cultural schism that replays itself even today, with the sustainability and environmental activists pushing towards environment (Eden) as the dominant value, while the industrial, commercial, and science and technology communities push towards technology and human progress as equally important (a human-built New Jerusalem).

What is interesting in this conflict between Jeffersonian agrarianism and a high-technology culture for the sustainable engineer, of course, is that the very idea of sustainable engineering can be seen in part as an effort to knit these two warring factions together, in a world that combines environmental quality with a high standard of living and a technologically sophisticated built environment. So in some ways, then, sustainable engineering itself is a response, hundreds of years later, to the railroad.

And what should we think of all this? Obviously, there are many themes, but most important, perhaps, as you look at your watch, reflecting patterns of time brought to you by the railroad, and as you work for a company shaped by railroad technology, in an environment and for nations who are what they are partially because of the railroad, with a mind shaped in part by the railroad . . . never, never underestimate what technology systems do.

5.3 TECHNOLOGY CLUSTERS

The constellation of social, economic, technological, cultural, theological, institutional, and policy patterns associated with a core technology is by no means unique to railroads. Indeed, the railroad case study just discussed is only one example of what economic historians call "long waves" of innovation. Sometimes referred to as **Kondratiev waves** after the Russian economist who played a major role in developing the relevant theory, each wave consists of a core technology, accompanied by coupled cultural, institutional, and economic changes (Figure 5.5).

As with railroads, none of the core technologies develop alone; rather, they all support "technology clusters." Automobiles, for example, support the development of petroleum-based infrastructures and the development of new information networks. For a mass consumption society such as the automobile helps create, information evolves into a mass culture phenomenon: radio, large circulation newspapers, television. Railroads and steam technology powered a wave from about 1840 to 1890; steel, heavy engineering,

Kondratiev Wave	Dates of Wave (approx.)	Core Industrial Organization	Time Construct	Capital and Finance	Critical Infra-structure	Consumption
Early Industrial Revolution (textiles, water)	1780–1840	Factory	"Day at work" model; work and home separated in time	(UK) capital from landowners and aristocracy	Water, canals	Mainly by hereditary elite, otherwise subsistence
Rail, steam, and coal	1850–1895	Joint stock company (railroads)	Coordinated regional times by firm (rail)	Development of modern finance (RR)	Local and national rail (coupled to telegraph)	Capitalists, some trickle down
Steel, heavy engineering, electricity	1895–1930	National firm, professional management (Taylorism)	Global time structure; mfg time models	Modern corporate finance	Rail, ocean shipping, roads	Capitalists, evolution of labor as class
Automobile, petroleum, mass culture industry	1930–1970	Mass production system	Time as integrating function for work and life	Mass credit, consumption financing (GMAC)	Roads, increasingly air; globalized infrastructure	Egalitarian consumption, mass consumers in core
Information, communication technology	1970–2000	Network-centric firm	Asynchronous and work time 24/7 as production globalizes	Finance as ICT; complex risk management	Networks	Rise of meritocratic elite consuming information, some trickle down
Nano, bio, robotics, ICT, cognitive science	2000–	Virtual, rapid adaptation, networks	Asynchronous, time as order of events	Global demat-erialized financial networks	Continued conversion of physical into global info networks	Meritocratic elite (shift from material/ financial to quality of life consumption)

FIGURE 5.5 Technological Change as Cultural Change. This figure shows the major Kondratiev waves since the Industrial Revolution began. Dates are approximate, and sometimes it is difficult to separate out the core technology from the supporting technologies. For example, are railroads the core of their wave, or is the steam engine? For this reason, it is best to understand these waves as new stable system states, representing a shift not just to new technologies, but new economic and social states as well. Every technology is itself only a part of a more complex system, and although identifying core technologies is useful for some purposes, the underlying unity among the technological, institutional, social, economic, and cultural systems cannot be forgotten.

and electricity, from about 1890 to 1930; the automobile, petroleum, and aircraft from about 1930 to 1990; the information cluster with its computerization of the economy, from about 1990 to the present. Although the dates are somewhat imprecise, the general idea of clusters of technology—which, it cannot be emphasized enough, always carry with them institutional, organizational, economic, cultural, and political changes—is a useful one. Each wave tends to generate characteristic institutional structures, especially within private firms. Thus, specialized professional managerial systems and associated "Taylorism" industrial efficiency techniques characterized the heavy industry cluster, while a far more networked, flexible structure began to evolve during the information cluster (characterized by, for example, telework and virtual offices).

Similarly, each wave generates changes in financial instruments, generally creating more complex, more flexible, and more rapid flows of credit and cash. Perhaps the most interesting example here is the coevolution of the automobile and mass consumption technologies, and the development of individual credit. Few if any individuals could save all the money required to buy a car, but they could make monthly payments on extensions of credit to buy a car; the white goods retailers, meanwhile, developed "lay away" plans that served the same function: enabling consumers to purchase expensive items for which they did not have ready cash. The information cluster evolves credit to be online, opening up finance and thus credit and consumption to global markets, with far more variety than many people could explore locally. The next phase, where credit is not just provided at the point of sale in a brick and mortar store, but provided over the Net for "B2C" (business to consumer) companies such as Amazon.com, has already become common; the integration of credit into cell phones, in both developed and developing country networks, is rapidly becoming so.

Our sense of time has also remained a function of technology: the information age, first with fax, then with email, instant messaging, and other technologies, has dramatically accelerated the sense of time fragmentation and time cycle. Many teachers, for example, lament the short attention span of students, which is in large part a direct result of their access to information technology and services such as multitasking, high-impact short video presentations on the Internet, and television.

So, if we as sustainable engineers create new technologies, it is also true that we ourselves are a product of previous technologies, and that our creations will shape the way future generations think, perceive, and interact with their increasingly virtual environment. It's not intuitive, but this historical perspective suggests that if you're designing a new information product or social networking service, you're also designing the way future generations will perceive their world (see Box 5.4).

BOX 5.4 Perception and Technology

Observations about how technological systems are perceived are common (the basis of much marketing, for example). Far less appreciated is how technology shifts personal and cultural perception. This is not just true of information and communication technologies, which might be expected to act as perceptual filters, but also of other types of technology. Consider, for example, the railroad's effect on the way people viewed their world. Writing in 1843 as railroad systems in Europe began their explosive growth, for example, the German Romantic poet Heinrich Heine powerfully captured the disorientation of traditional psychological framings of time and space that his generation experienced when confronted by the new technology:

> What changes must now occur, in our way of looking at things, in our notions! Even the elementary concepts of time and space have begun to vacillate. Space is killed by the railways, and we are left with time alone. . . . Now you can travel to Orleans in four and a half hours, and it takes no longer to get to Rouen. Just imagine what will happen when the lines to Belgium and Germany are completed and connected up with their railways! I feel as if the mountains and forests of all countries were advancing on Paris. Even now, I can smell the German linden trees; the North Sea's breakers are rolling against my door.[3]

The automobile had the same effect on many people. Especially for many urban populations in the early 1900s, the automobile represented a major means by which they could perceive "nature." Thus, for example, the Bronx River Parkway, built in 1925 to connect New York City and Westchester County, was designed by Fredrick Law Olmsted, the same landscape architect who designed Central Park, to be a "scenic utopia" reflecting an "authentic" American landscape. As an acute observer of urban design writes, "nature became simultaneously more distant (framed by the window of a moving car), more accessible (through greater public contact with remote areas), and at the same time more individualized as an aesthetic experience."[4] To this end, in fact, the Parkway was planted with thirty thousand carefully selected trees to present an image—to the person in an automobile—of what "natural America" was supposed to have looked like. To perceive that Parkway from an automobile was to perceive "real nature." This example is particularly interesting, of course, because to so many people "nature" is not a matter of perception, but something unchangeable, almost sacred, that exists on its own, independent of humans.

The mediation of human perception through technology is a phenomenon of long-standing. Medieval clerics knew what the moon looked like, because Church

---

[3]Quoted in Wolfgang Schivelbusch, *The Railway Journey: The Industrialization of Time and Space in the 19th Century* (Los Angeles: University of California Press, 1977), p. 37.

[4]See M. Gandy, *Concrete and Clay* (Cambridge: MIT Press, 2002), p. 122.

authorities knew what God had meant the moon to look like; the technology of the telescope did not just allow scientists to observe the moon, but overthrew centuries of theology and belief. But modern experience is perhaps more familiar, and it reinforces the critical role of technology in human perception. The shorter attention span characteristic of many young people today, especially the "digital natives," is in large part a function of their exposure to rapidly changing, high-impact, media, especially modern television and Internet technologies. And think of how "nature" changes as it is seen through the window of the television and Internet, more distant and yet, thanks to the power of nature photography, appearing to be even closer and more accessible, and more emotionally evocative. Of course, nothing is more artificial and contrived than nature programs on television and cable, but it is that image of nature, a product of the technology and mass consumption culture, that for many people has become "real." And, of course, the impacts that virtual realities will have on how people perceive the "real world" are just beginning to be considered, and are certainly not yet understood.

Technology, in short, is not just the way in which natural systems are affected by humans. It is also the lens through which the world and the natural systems that compose it are perceived and understood by humans: we can only perceive what our technology allows us to. Without computers and computer models, for example, it would be impossible to perceive global climate change. To paraphrase the movie line, "You want the truth?! You can't even perceive the truth!" The reason is that, in a high-technology world, just like quality of life, or number of people that can be supported, what can be seen, and understood, is a function of the available technology.

Exercises

5.1 **a.** You are the chief technical officer of a startup firm that has developed a nanotechnology-based system that enables inexpensive and energy-efficient capture of CO_2 from the ambient atmosphere. Using the categories of Figure 5.5 as a guide, prepare a two-page memorandum to your chief executive officer identifying some of the broader implications of your technology.

b. Write a follow-up memorandum discussing how you might present your technology to the public in such a way as to reduce potential backlash from environmentalists and the public.

5.2 What are some of the military implications of the ambient atmosphere carbon capture technology identified in Exercise 5.1?

5.3 Is the idea that one's genetic make-up affects the way one behaves in social situations an example of technological determinism? Why do you suppose it is such a widely held opinion? What are some of the likely consequences of a government and its policy makers maintaining this belief?

5.4 What are the different ways in which the automobile has started to domi-

nate "social space" in contemporary cities? To what extent has it altered the way humans interact with each other (a) in your own context; and (b) in a different cultural context that you may be aware of or familiar with?

5.5 Why do you think technology clusters have institutional, social, and cultural effects, rather than just economic impacts?

Annotated Bibliography

1. Although there are many specialized studies of railroad history, it is hard to find good general treatments. One of the better and more interesting is W. Schivelbusch's 1977 study, *The Railway Journey: The Industrialization of Time and Space in the 19th Century* (Berkeley: University of California Press). For those seeking a different perspective on what the railroads could do, W. Cronon's masterful 1991 tome, *Nature's Metropolis: Chicago and the Great West* (New York: W. W. Norton and Company) is hard to beat; it serves as an excellent integrated treatment of environment, urbanization, history, and technological evolution. More generally still, the question of technology, history, and natural systems is covered quite well in A. Grubler's 1998 book, *Technology and Global Change* (Cambridge: Cambridge University Press).
2. An excellent if occasionally technical text on Kondratiev waves is C. Freeman and F. Louca, 2001, *As Time Goes By: From the Industrial Revolutions to the Information Revolution* (Oxford: Oxford University Press). A broader treatment of the history of trade over the past two millennia is a nice little book by R. Findlay and K. H. O'Rourke, published by the Princeton Press in 2007, *Power and Plenty: Trade, War, and the World Economy in the Second Millennium;* interested readers might also like P. Hugill's 1993 book, *World Trade Since 1431* (Baltimore: Johns Hopkins Press).
3. It is often hard for those who have not studied technology to understand how human and technological components fuse to create emergent behaviors and cognitive networks that are far more powerful—and far more dependent on technology—than would otherwise be the case. The example of navigation systems used by E. Hutchins in his 1995 book, *Cognition in the Wild* (Cambridge: MIT Press) is excellent in this regard. A. Clark's 2003 book, *Natural Born Cyborgs* (Oxford: Oxford University Press) presents a more philosophic perspective on how humans have always been integrated with their technologies, especially at the cognitive level.
4. For those seeking to understand how technology and culture interact, especially in the case of the United States, three books recommend themselves: J. Nye's 2003 *America as Second Creation: Technology and Narratives of New Beginning* (Cambridge: MIT Press); his 1994 book, *American Technological Sublime* (Cambridge: MIT Press); and the book of the doyen of the domain, Leo Marx's 1964 *The Machine in the Garden: Technology and the Pastoral Ideal in America* (Oxford: Oxford University Press). These are not difficult books,

but they are very useful for a deep understanding of technology. Two good books that cover the same topic, but in a European Enlightenment rather than an American framework, and with more emphasis on the easy-to-overlook but critical relationship between Christianity and technology in the West, are M. H. Abrams' 1971 work, *Natural Supernaturalism: Tradition and Revolution in Romantic Literature* (New York: W. W. Norton & Company), and M. H. Nicolson's 1959 study, republished in 1997, *Mountain Gloom and Mountain Glory: The Development of the Aesthetics of the Infinite* (Ithaca, NY: Cornell University Press).

CHAPTER

6 Characteristics of Technology

6.1 INTRODUCTION

It is almost pathetically oversimplistic to have one chapter dealing with characteristics of technology. Necessarily, the content of such a chapter will not be comprehensive and will reflect a difficult process of triage of many important discussions. That said, there are some important insights and themes that are worth mentioning, and with which the sustainable engineer should be at least acquainted. Accordingly, this chapter will begin by looking at a fairly specific and common activity, product design, in order to provide some grounding in how artifacts come about, and end by proposing a way to conceptualize technology systems, including both artifacts and the goals they are intended to meet. In between, we can hope for a degree of enlightenment.

6.2 PRODUCT DESIGN

One way to think about technology is to consider the general challenge posed by product design. To begin with, no product can be made without using materials, and usually investing significant energy in the changes in state and shape required for any product. Manufacturing any product, therefore, has at least potentially negative environmental, and thus sustainability, implications. Conversely, however, products are not produced, at least for long, without consumer demand—whether from individuals, states, institutions, or other firms—and thus they have some implied positive social value. From the very beginning, then, any product represents at least a potential conflict between the two themes

of sustainability. For the sustainable engineer, this implies that neither a purely economic nor purely environmental perspective is likely to be satisfactory, and that good design and application of sustainable engineering skills will likely focus in many cases on reducing the environmental footprint of a product over its lifecycle while not jeopardizing its economic or social value.

Product design activities are also highly constrained from the beginning. Products must be manufacturable at a competitive price in a competitive time frame with characteristics that appeal to customers, and must intersect easily with all important infrastructures and coupled technologies. (For example, if you build an electrical product for global markets, it will need different plugs for different countries.) Moreover, products must comply with all applicable laws and regulations, not necessarily an easy trick in a world where products may be sold into many markets with different regulatory and safety regimes. Environmental regulations are especially tricky: they can change frequently, they differ markedly between jurisdictions (the European Union, for example, is prone to banning materials of concern), and they are frequently manipulated by domestic firms to create subtle trade barriers. For the sustainable engineer, these constraints, and many more, are an integral part of the difficult and subtle problem-solving activity called "engineering design." The degrees of freedom actually available in any design are often very limited, so the ability to optimize on only one or a small set of values—environmental performance, sustainability—is as well.

There are many different design approaches used across industry. The tools that a particular engineer or design team might choose will reflect the type of product, the experience of the design team, firm and sector habits, and the like. Many of these, such as the House of Quality tool, incorporate customer demands and desires into the process in a way that facilitates their quantification, and identifies potential conflicts and synergies. For example, when designing an automobile, a well-soundproofed car may signal quality, but it is also heavier, and thus has lower gas mileage, than alternative designs. Conversely, for some consumers an aluminum vehicle may carry a high-tech cachet; it also lightweights the vehicle, thus enhancing performance, which such consumers also find attractive.[1]

Some idea of the complexity of the design process can be obtained by considering various modules of one popular process known as "Design for 'X,'" where "X" is a particular desirable attribute of the product. Thus, one may have the following:

Design for Assembly (DFA): Is the product easy to assemble with inherent high quality (that is, does the assembly process itself reduce errors

[1]More information on the House of Quality and other techniques can be found in Thomas Graedel and Braden Allenby, *Industrial Ecology and Sustainable Engineering* (Upper Saddle River, NJ: Prentice Hall, 2010).

and shoddy construction)? This is a very important consideration for controlling costs, because it is much less expensive to have high-quality product runs in the first place, rather than to fix bad product coming off the line before shipment. From a sustainable engineering perspective, less bad product equals less unnecessary waste and less unnecessary cost.

Design for Disassembly (DFD): Can the product be easily disassembled (using snaps, for example, rather than screws) for subsequent recycling or reuse of components or materials? This is tricky in part because one must make assumptions about the technologies that will be used at product end-of-life: if one simply sends the stuff to a dump or to a hammermill for smashing before recovery of metal content in a secondary smelter, a lot of fancy design is obviously unnecessary. Moreover, some products have unpredictably long lives: the American B-52 bomber, for example, was first bid in 1946, test flown in 1952, and entered active service in 1955, and is still flying. Obviously, any design assumptions made when the first design work was ongoing in the late 1940s would be completely obsolete today (and the culture in the 1940s would have been completely befuddled by the idea of design for disassembly as well). From the sustainable engineering perspective, DFD requires understanding the systemic context of the product and the technologies involved at each stage of the process.

Design for Environment (DFE): Is the product designed to be environmentally sound, with proper selection of materials and characteristics such as long life and energy efficiency that reduce its environmental footprint? This module may often rely on tools such as Life Cycle Assessment (LCA) to investigate design options, assuming there is the time and resources for using them. Again, long-lived products may be problematic, because information on environmental impacts and valuation of different types of impact shift over time. For example, in the 1970s cleaning up manufacturing dump sites and emissions was a priority; today, climate change is the dominant priority. Designing to avoid use of toxics in manufacturing in the 1980s—which involved, for example, substituting water-based solvent systems for chlorinated hydrocarbon solvents in piece-part cleaning operations—increased energy use to reduce workplace hazard. Today, options that increased the "carbon footprint" are much more highly disfavored, so such a change might not pass DFE muster.

Design for Manufacturability (DFM): Is the product easy to manufacture? Does it facilitate environmentally and socially responsible manufacturing by, for example, having components that can be cleaned without the use of toxic materials? Does it reduce production of products

that fail quality standards, thus reducing manufacturing waste? From a sustainable engineering perspective, much of the environmental and social impacts of manufacturing can be ameliorated not just by proper manufacturing line design, but through good product design as well.

Design for Regulatory Compliance (DFRC): Is the product compliant with all regulations and laws in its relevant markets (which for many products means global compliance)? From a forward-looking perspective, is the product easy to redesign to meet laws and regulations that are likely to come into effect during its relevant lifecycle? Depending on how broadly we want to define this "X," we might include product liability, safety laws, applicable standards (even if not legally required), and other constraints in this module. This is an area where sustainable engineering can significantly reduce subsequent costs: A flexible product design (perhaps modular, for example) can greatly reduce costs by enabling agile adjustment to regulatory change, as well as facilitating incremental product redesign and improvement going forward.

Design for Reliability (DFR): Given the range of environments in which the product is likely to be used, and how it will be used, is it reliable and safe? For example, an electronic device used where there's a high risk of static electricity must be designed to avoid failure when exposed to static discharge. From a sustainable engineering perspective, DFR can extend product life, which is generally, although not always, a good thing. Note that in such cases sustainable engineering aligns environmental values with social values: Customers like reliable products, and the environment appreciates the reduction in waste that product reliability can generate.

Design for Sustainability (DFS): Is the product sustainable? This is obviously a fairly new category, and as of yet tends not to be well defined, but it might include social, economic, and environmental factors that are not picked up in other modules. The danger here is that identified in the chapter on sustainability: DFS simply becomes a mechanism by which particular ideologies are expressed.

Design for Service and Maintenance (DFSM): Is the product designed to minimize the installation, servicing, and maintenance operations that are necessary? For example, a sustainable engineer might design a weapons system to rely heavily on modular design to reduce the sophistication of field maintenance required, given the unsettled conditions that often characterize combat environments. Like many design categories, this one contains potential conflicts: If a product is designed so that servicing and maintenance are minimized, this can also mean that when the product breaks down, it is cheaper for the consumer to buy a new one than to get the old one fixed, potentially adding unnecessarily to the waste stream.

Design for Testability (DFT): Is the product easy to test? This is a particularly important characteristic for very complex products, such as electronic chips or computers, because it may be difficult, even conceptually, to know how to robustly test performance. A new high-end chip, for example, may perform millions of functions, and without prior design consideration may be almost impossible to test. Similarly, complicated software, especially if it is used in an "open" environment, such as home computers attached to the Net, is notoriously difficult to test because of the complexity of the system within which the software functions.

Many more modules might be used, depending in part on the particular product and the specific design issues associated with it. Moreover, these sorts of product design processes require expertise in very different areas, which is one reason modern product design is usually done by teams. For example, the law department would most likely provide input on applicable regulations and potential liabilities that should be addressed, while the environment, health, and safety group (or corporate equivalent) would provide the technical expertise on environmental implications of product design and manufacture options. Even among the engineering team members, it is likely that some would be expert in design of manufacturing processes, some in product design, some in material science, and, increasingly, some expert in lifecycle assessment of product impact and performance. Note, however, that design is always a balancing process, with many different constraints and objectives jockeying for position; what emerges must reflect real-world conditions, rather than wishful thinking. Thus, for example, sustainability activists are free to demand that all toxics be eliminated from a complex product, but it is a rare design indeed where the engineers can do that and at the same time meet all other relevant considerations. Similarly, it is often the case that activist groups do not reflect the characteristics of all consumer segments in the marketplace; firms can lead, but at least in a free market they must also be responsive to consumer demand.

So what is the process that a product might go through before it actually gets to the market? Figure 6.1 maps out a typical route, with decision points, usually called "gates," noted as numbers. At each gate, either a formal or informal review by the team responsible for the product, which might include not just the design team members, but people from marketing, sourcing, logistics, external affairs (an organization responsible for both representing the firm to external audiences and informing the firm about looming developments, such as material bans, that might affect the product), is held. The continued commercial viability of the product is assessed, given the product characteristics that are increasingly locked in as the design proceeds. The reason for such continued assessments is clear: The earlier an uncompetitive product is identified

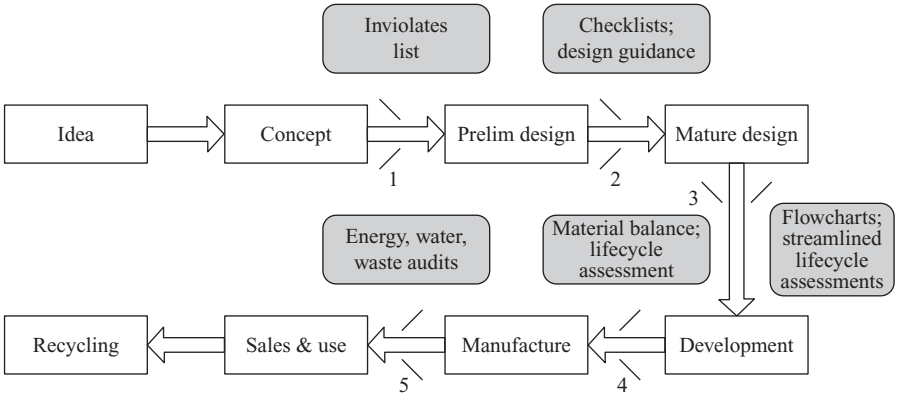


FIGURE 6.1 Environment and Sustainability in the Product Realization Process. This is an overview of a generic product realization process. In addition, the boxes outside the flow process show possible environmental and sustainability assessments, from lists of materials or particular products that should not be considered (the "inviolates list"), to gates, indicated by numbers, where various assessments might help inform the "go/no go" decision.

(Based on: Graedel, T. E. and B. R. Allenby. *Industrial Ecology*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, © 2003, p. 235. Reprinted by permission of Pearson Education, Inc.)

and killed, the less time, money, and resources are wasted. Thus, each gate not only informs the next stage of the design process, but is an important governance mechanism assuring robust design. From a sustainable engineering perspective, these gates become even more important, as they are existing mechanisms for managing the ever increasing demands of stakeholders and NGOs on firms, their products, and their activities around the world. A good gate process can be an important part of institutionalizing sustainable engineering. Conversely, an inadequate one—say, one that ignores the requirements of important markets, or of powerful activist stakeholders, or fails to identify important sustainability issues so that they may be addressed early in the design stage—can make a design less efficient, more expensive, and perhaps inadequate.

6.3 BEHAVIOR OF TECHNOLOGICAL SYSTEMS

Despite the fact that technology is a very complicated phenomenon, there are some regularities worth noting. Most obviously, virtually all technologies, at whatever scale, have lifecycles that display the same general pattern. Thus, as Figure 6.2 illustrates, technologies generally have a research and development phase, which includes a number of activities: basic research, applied research, invention of the actual technology, and an innovation step where the invention is modified to be economically and commercially viable, and the necessary institutional changes or developments to support it as a commercial operation

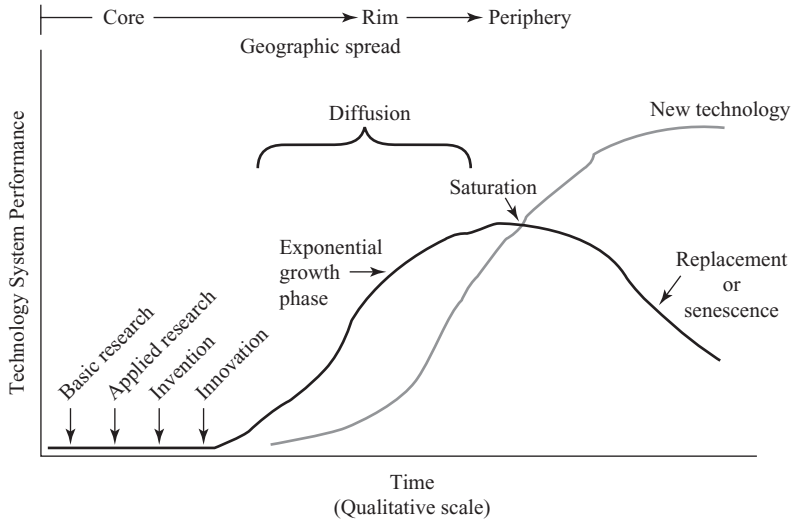


FIGURE 6.2 Idealized Technology Lifecycle Model. This figure illustrates an idealized schematic of the lifecycle of a typical technology, from development, to diffusion to rim and then to periphery markets, and then to eventual replacement by a new technology. How long each technology remains viable varies significantly depending on its characteristics: Some infrastructure technologies have lifetimes of many decades, while electronic technologies may have lifetimes measured in months.

(Based on: Graedel, T. E. and B. R. Allenby. *Industrial Ecology and Sustainable Engineering*. 3rd ed. Upper Saddle River, NJ: Prentice Hall, © 2010, p. 69. Reprinted by permission of Pearson Education, Inc.)

occur. Although Figure 6.2 shows these steps as following each other, in most cases they are overlapping, and sometimes one or more may be lacking. Many technologies, for example, are introduced and become commercially successful well before the underlying basic research is complete.

At the end of the development process, a commercially viable technology enters a "takeoff" period, where it rapidly diffuses into the economy. It should be noted that this does not mean that the technology does not continue to evolve; the Japanese are famous for their incremental improvement processes, which is part of what has made their manufactured products consistently among the best and most competitive in the world. Innovation, in other words, may continue throughout the lifecycle of the technology. Moreover, incremental improvements are not necessarily either rapid or trivial. Figure 6.3 illustrates the implementation process required for a typical material substitution design change. The core process moves from concept to implementation, and is driven by economic analysis of the proposed changes, but a number of different activities and analyses feed into that process. Some, such as prototyping and hardware demo, can be accelerated in some case through use of computer tools. Others, such as durability testing of production components, are done late in the implementation process. In some cases, the substitution can be

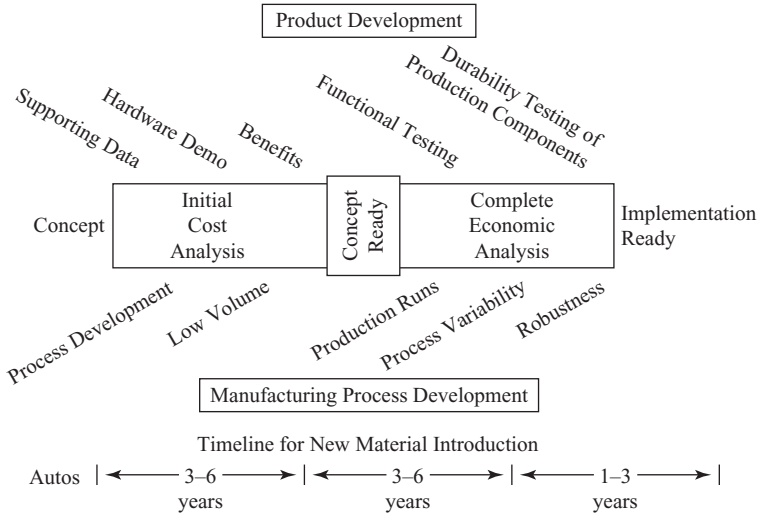


FIGURE 6.3 Technology Implementation Process. This figure illustrates the schematic road map for making a material change to an existing product; both the product and the manufacturing process need to be thoroughly vetted to ensure safe, reliable manufacturing and product performance. The process can require significant time, even in today's environment where at least some steps can be augmented by computational modeling and machine-based design and expert systems.

(Based on: presentation by John Allison, Ford Research Laboratory, to NRC Materials Advisory Board study group, March 13, 1998.)

done relatively quickly, but in others it is a much slower process, as the numbers suggested for the automobile industry indicate. In general, virtualization software and enhanced computer programming have been able to trim time off this cycle, but it is still frequently a matter of years, especially where, as with automobiles, safety issues are critical as well. In fact, recent research shows that on average countries have adopted technologies, such as cars, trucks, personal computers, telephones, or magnetic resonance images, over forty-five years after their invention. This number, however, hides substantial variations depending on technology and country, and a general trend for more rapid adoption of newer technologies as compared to older ones, and a shorter time to obsolescence, even for core technologies such as energy or transport technology systems. Such trends are to be anticipated given increasingly rapid rates of technological evolution.

Moreover, this is only the technical aspect of material substitution. The more a manufacturer is responsible for products (and thus materials) over the lifecycle, the longer it can take. For example, from a sustainable engineering viewpoint, any product that uses any material from a nation or region involved in conflict increasingly may engage human rights issues, and thus require additional analysis and care. The classic example to date is the attacks on the electronics industry by human rights activists for their use of tantalum from

"coltan," the colloquial name for an ore combining columbite and tantalite, mined in the Democratic Republic of the Congo, which allegedly supports the conflict in that area. Despite the fact that the DRC produces less than 1% of the world's tantalum, the public pressure generated by NGO campaigns has created a significant problem for firms using tantalum in mobile phones, DVD players, and other electronics equipment. Sustainable engineering would thus suggest designing products and selecting materials to avoid such issues in the first place, where possible. Where the particular material is critical, the sustainable engineer might want to explore alternatives, such as establishing a "chain of custody" process that ensures that purchases of the material in question are not used by terrorists or insurgents to fund conflict.

Returning to Figure 6.2, the takeoff period is typically marked by several different types of growth. On the one hand, the technology grows exponentially in its core markets, which are those areas where the technology is developed and conditions are optimal for its rapid adoption. The technology also spreads into rim and then into periphery markets, perhaps only when appropriate adaptations are made. Thus, for example, expensive multifunctional personal digital assistants are common in North America, Europe, and developed countries in Asia, but may more slowly penetrate rim markets such as the BRIC countries (Brazil, Russia, India, and China), while only second-hand units, or units containing less functionality and requiring less bandwidth, may penetrate peripheral markets in Africa or poorer parts of Asia. In addition, both business models and product may need to be modified as they migrate from core to rim to reflect different use and disposal environments. Thus, for example, the business model for mobile phones in core markets usually relies on monthly payments after use, where the service provider is essentially providing credit for the user. Because of the poverty in periphery markets, this payment model is difficult to replicate in such areas, so providers have developed means, such as "pay-as-you-go" offerings, by which users can buy service as they need it. The speed with and degree to which advanced technologies become widely available are important factors for sustainable engineering; although it is true that technology spreads from the (usually more wealthy and educated) early adopters to others over time, there is continuing concern that the speed of technological evolution means that poorer areas of the world are left increasingly behind in terms of technological availability and competence, despite the occasional optimistic story such as the rapid adoption of mobile phones in many poorer countries. There are also indications that patterns are changing: Multifunctional digital assistants may be widely adopted by world elites and less so by world proletariat, rather than spreading through the historical core/rim/periphery pattern.

So the technology spreads by both rapid adoption in the core and by extension to noncore socioeconomic cohorts or geographic regions. In both cases,

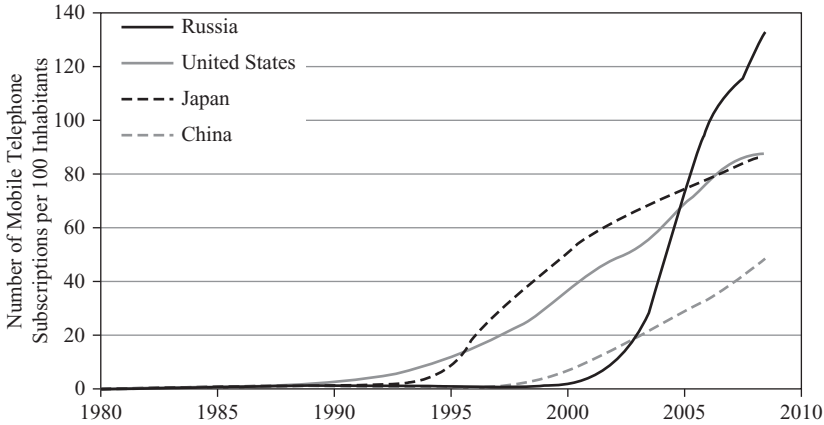


FIGURE 6.4 Mobile Telephone Subscriptions per 100 Inhabitants. This graph shows the growth of mobile telephone technology in four major markets. Note that the general shape of the curves are similar, even as details differ depending on many factors (e.g., Japan as a small country with concentrated population centers had a more rapid growth phase compared to the United States, where more of the population was geographically diffused across large distances). China also shows a very rapid uptake of the technology because the competing technology, landline telephony, is less available and more expensive than, say, in the United States, where uptake of mobile telephone technology is relatively slower.

(Data from: United Nations Statistics Division. "Mobile Cellular Telephone Subscriptions Per 100 Inhabitants." United Nations. http://data.un.org/Data.aspx?q=mobile&d=ITU&f=ind1Code:I911: accessed June 4, 2010.)

however, the technology at some point achieves saturation, the point of maximum demand and market penetration. As saturation is neared, growth slows, and hence one gets the "S curve" familiar to virtually all students of complex systems. Although the specifics of each curve may differ, it is remarkable how the generic shape remains similar. This can be seen in Figure 6.4, which illustrates the growth of mobile phones in four major markets: Japan, the United States, Russia, and China. Such charts need to be understood carefully, however; for example, the functionality available in Japan over mobile telephones is significantly greater than that available in the United States, in part because more Americans tend to access Internet over their computers, as compared to Japan, where much Internet functionality is accessed over mobile telephone.

Compare, for example, Figure 6.5, showing the growth in ownership of Internet-enabled personal computers, with Figure 6.4. The most obvious difference is that the technology is less mature. The shape of the curves for Japan and the United States, show that their markets are maturing, but the shape of the curve for Russia indicates that it is still in its exponential growth phase while the shape for China indicates that it may be still in the last part of the takeoff phase. If, for example, you were a sustainable engineer working for a firm making Internet-enabled machines—or, perhaps, planning a substitute for them, such as an enhanced personal digital assistant designed for the

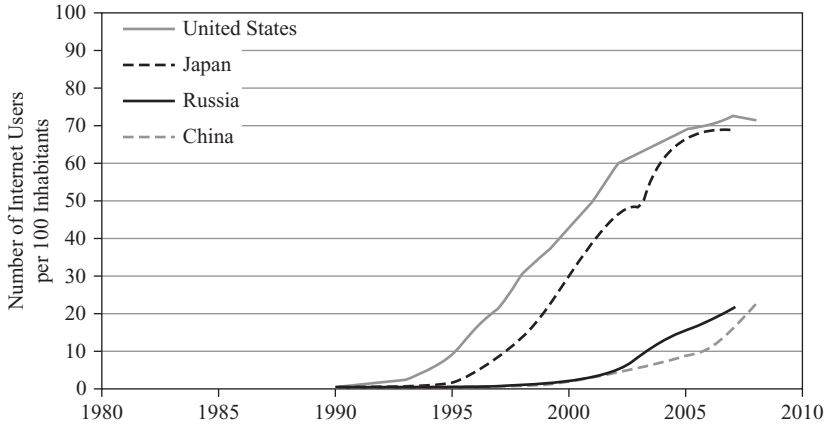


FIGURE 6.5 Estimated Internet Users per 100 Inhabitants. This figure illustrates the growth of Internet use. Note that this adoption appears to be leveling out at approximately 70% of the population. This doesn't necessarily translate into sales levels, of course, which depend on many other factors.

(Data from: United Nations Statistics Division. "Estimated Internet Users Per 100 Inhabitants." United Nations. http://data.un.org/Data.aspx?q=internet&d=ITU&f=ind1Code:I99: accessed June 4, 2010.)

Chinese market—such observations are of serious interest from both a product design and a competition perspective. The second obvious observation arising from comparison of these two figures is that, for whatever reason, society apparently regards mobile phones as more necessary than Internet-enabled personal computers. This is apparent because the penetration of mobile units begins to level off at around the 90% household mark, whereas for the computers both Japanese and U.S. penetration appears to be leveling off at around the 70% level.

In addition, it is important to remember that none of these technologies arise in a vacuum. A mobile telephone does not operate absent a viable network, and the services that can be platformed on it are determined by the bandwidth and sophistication of that network. Moreover, the institutional and cultural dimensions of a particular market are also critical: many mobile phones in Africa are being used as minibanks, able to transfer sums around a society where lack of financial infrastructure and physical insecurity makes cash transactions difficult if not impossible. Such services are not generally offered in developed countries, where the need is not nearly as great. Similarly, an Internet-enabled personal computer is just a stand-alone computer in the absence of an Internet connection, and that connection has to work with the software and hardware with which the computer is fitted. In a developed country such as Japan, Korea, or Finland, you may be able to assume the availability of whatever suite of coupled technologies is necessary; in many areas of Latin America or Africa, that may not be a viable assumption.

Saturation, the peak of the S curve shown in Figure 6.2, is an interesting point. With some technologies, incremental improvements continue but the core technology remains current, so that the S curve levels out, but does not show an immediate replacement technology. Such patterns are especially true with technology systems, such as automobiles or the railroad, which are foundational for a Kondratiev wave. Such technologies are **locked in**, in that they are so tightly coupled to other systems (such as petroleum and road infrastructures) that any major change to the technology faces almost insurmountable odds. (That's why Kondratiev waves are so revolutionary.) With other technologies, such as electronic components, the stability may lie in the interface between functional units, while the functional units themselves might change radically. Thus, for example, the continuing increase in computer chip power mentioned in the last chapter may over time be entirely revolutionary, but because the interface between the chip and the rest of the technology platform is relatively stable, the evolutionary process is relatively free from lock-in in that sense. In other senses, of course, design is locked in: Manufacturing equipment assumes a particular chip production technology, for example, and the heavy institutional and intellectual investment in silicon as opposed to other alternatives is a subtle form of lock-in on its own.

Two other important factors may affect the saturation point, especially with rapidly evolving technologies such as electronics. First, the categories used in such analyses tend to reflect products, and product categories in sectors such as electronics are not stable. For instance, the original digital cameras were one-for-one replacements for film cameras, but digital photography is increasingly becoming an add-on function of mobile phones. So simply tracking digital cameras provides an increasingly misleading indicator of the popularity of digital photography. Similarly, the function of accessing the Internet has been different in different countries: in Japan, for example, access has generally been through mobile phones; in the United States, through personal computers. Such differences ripple across technology systems in broad and interesting ways. The Japanese wireless infrastructure has had to be more advanced and generally has had to provide greater bandwidth than the United States because of differing demands, and mobile phone software for the two countries obviously has to be different to account for different cultural patterns. Similarly, portable music evolved from cheap transistors playing poor-quality AM radio, to the Walkman, then to the iPod. As evolution continues, the iPod itself becomes a part of the mobile phone technology complex (or is it the mobile phone becoming part of the iPod?). Meanwhile, the mobile phone enables completely new technologies, such as augmented reality (the ability to access Internet and other data sources and display them over the real world in real time). Such trajectories warn against becoming too doctrinaire about any aspect of technological change.

On a more prosaic level, companies and their engineers are frequently faced with the question of "cannibalization" at the saturation point. That is, part of what creates a saturation point is a new generation technology beginning to replace the older one, and in doing so creating conditions of rapid obsolescence. This can happen with specific products, such as computer chips, raising strategic questions about whether, and when, a firm should cannibalize its own product, even if it is still profitable, to prevent a competitor from doing it and thereby achieving market dominance. It can also happen broadly across sectors. Consider, for example, Figure 6.6, illustrating shifts in major U.S. transportation infrastructures. In some ways, of course, this is the same point made in the previous figures: you get the S curve growth pattern, then substitution of a new technology. It also, however, provides an important cautionary note to sustainable engineers: if you are interested in changing a technology that is tightly coupled to such fundamental infrastructures—transportation, energy, water, built environment in most areas—then the time cycle for change may not be a product

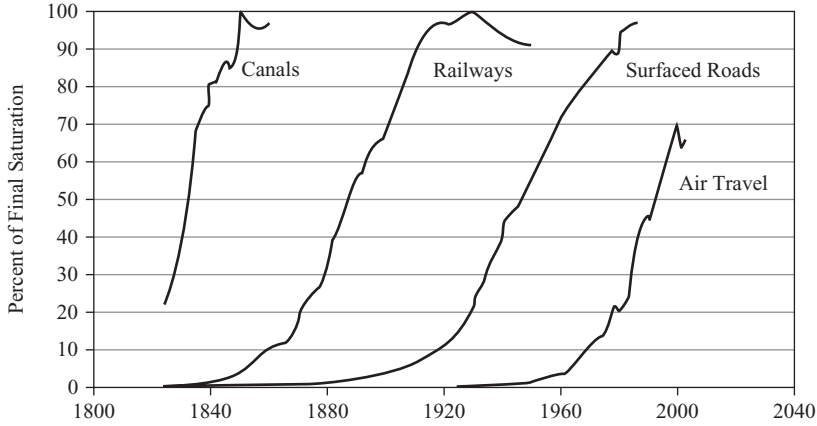


FIGURE 6.6 Major U.S. Transport Infrastructure. This graph illustrates the shifts between major transport infrastructure systems. Note how these shifts reflect the Kondratiev waves of technological change, as railroad and road infrastructure are necessary coupled technologies to railroad train and automobile mobility technologies. Note also the vertical axis, which is in terms of percent of eventual saturation level (roads and air travel have not yet reached this point); if another metric, such as miles of infrastructure, were used, the newer infrastructure would completely swamp the older one, giving a misleading idea of the importance of each infrastructure in its time. In addition, note the regularity in infrastructure substitution, with new technology becoming dominant approximately every five decades. Note: Canals, railways, and roads are measured in miles of transport infrastructure. Air travel is measured in revenue passenger miles.

(Adapted from: Grübler, A. *Technology and Global Change*. New York: © Cambridge University Press, 1998. Reprinted with permission; Additional data from: USCB, S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright, eds. *Historical Statistics of the United States on CD-ROM: Colonial Times to 1970*. New York: Cambridge University Press, 1997, Series D152-166; Various editions of the United States Statistical Abstract available from http://www.census.gov/compendia/statab/.)

of your technology, but of the infrastructure to which it is attached. Thus, for example, the time required to change to hydrogen vehicles is not a process that will be mediated by designing such vehicles but, among other things, the much more substantial time required to develop and implement a hydrogen energy infrastructure, whatever shape that might take in practice.

The periodicity of changes in infrastructure suggested by Figure 6.6 is, of course, reflected in the economic cycles that are part of the intellectual framework of Kondratiev waves. The railroad, steel and heavy engineering, and automobile waves, for example, lasted about 100 to 120 years, which is reflected in the rail and road infrastructure shifts.

Figure 6.6 also suggests another phenomenon that can occur at the level of specific technologies or technology systems. Sometimes, rather than an incremental substitution of one technology for another, we get a technology that fulfills many of the functions of an older one, but does so in a much more effective or different manner, creating significant technological and economic perturbations along the way. Examples of such so-called **disruptive technologies** might include digital cameras that, in a relatively short time, displaced film cameras in a shift that left some traditional firms, such as Kodak and Polaroid, struggling to adapt. A looming disruptive technology might be electric cars, although note that the infrastructure required by digital cameras (computers for processing, Internet for distributing, and the like) already existed, but the smart grid required for electric cars has yet to evolve. For the sustainable engineer, the relevant questions about such disruptive technologies would be familiar: Does the technology provide environmental, economic, and socially preferable functionality? Can it be designed so that it is not just technically elegant, but commercially successful? From a business perspective, the most important question is, When, and how, can a firm be transitioned from an older technology to the disruptive one? It should be noted that such transitions usually occur by the firm producing the older technology simply being displaced by newer firms, suggesting that institutional barriers and complexities often preclude the necessary changes within firms.

But the Kondratiev waves are not the only important long-term trend that economic historians have observed. We close this section with Figure 6.7, which illustrates sector employment in the United States from 1840 to 2000. The pattern there, which shows a significant shift away from agriculture initially to manufacturing, then a dramatic growth in services as the economy modernizes, is replicated in most developed countries. Even countries like Japan, where manufacturing remains an important part of the economy, and a national competency, show similar patterns. It must not be forgotten, however, that the economy has grown significantly over this period, so that it is more accurate to say that manufacturing as a percent of the economy, either by employment or by revenues, is down, but in absolute terms remains important: In 2007, the United States remained the largest manufacturer in the world.

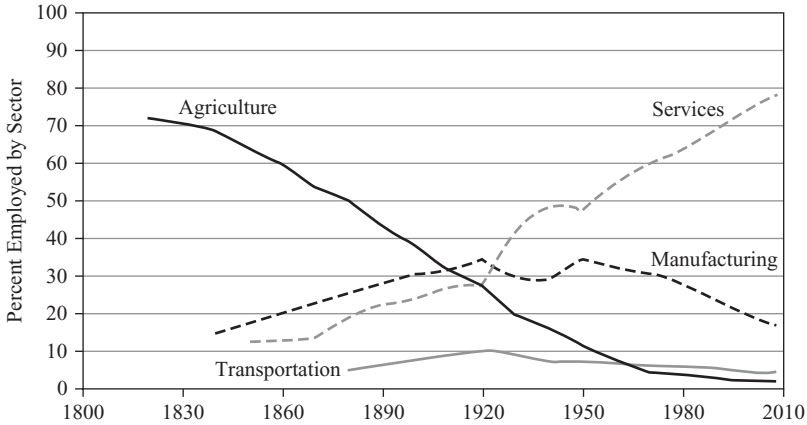


FIGURE 6.7 United States Sector Employment. Over time, services have grown to dominate employment (and other metrics, such as percent of GDP). This has occurred not because the manufacturing and agricultural sectors have shrunk, but because most economic growth has occurred in services. It should also be noted that it is increasingly difficult to draw clear lines between manufacturing, services, transportation, and agriculture (is integrated pest management a service or agricultural activity?).

(Data (1820–1920) from: USCB, S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright, eds. *Historical Statistics of the United States on CD-ROM: Colonial Times to 1970*. New York: Cambridge University Press, 1997, Series D152-166; Data (1930–2008) from: U.S. Department of Commerce Bureau of Economic Analysis. "National Income and Product Accounts Table: Table 6.8. Persons Engaged in Production by Industry." http://www.bea.gov/national/nipaweb/SelectTable.asp: accessed June 2, 2010.)

Nonetheless, the shift to services is important for several reasons. First, it illustrates in a very salient way the importance of intellectual capital in modern economies. Consider, for example, agriculture. Advanced agricultural technologies, such as integrated pest management (IPM), low-till, and genetically engineered products (GMOs), have significantly augmented agricultural productivity. Although agriculture is still a process of putting seeds into the ground and getting them to grow, the intellectual capital input to that process has changed it enormously even over this century, a trend that is accelerating dramatically with GMO technology. If electric cars and highly variable energy production technologies such as solar and wind are going to become important parts of the energy picture, they will require a smart grid—or, in other words, a grid that is as much about service and concomitant intellectual capital as it is physical infrastructure. A car remains a mechanical vehicle, but the amount of information management in the modern car, and increasingly built into information support systems such as satellite radio and traffic management, makes transportation a far more information dense activity than before, and again blurs the line between artifact, artifact in use, and information service.

The shift to services is also important because some people think that a service economy can provide equivalent quality of life, with less environmental

footprint, than a product- and consumption-oriented economy. This is best taken as a working hypothesis, rather than as a reality, certainly until more data and experience are accumulated. After all, human life and institutions require a significant flow of energy and materials, and all services are platformed on underlying physical technologies of one sort or another. Nonetheless, anecdotal support for the hypothesis exists, and it may be an important element of sustainable engineering in the future.

To some extent, this section demonstrates the deep challenge of sustainable engineering. Not only must the sustainable engineer be good at the traditional aspects of engineering, but this person must be able to understand, at least in theory, the function of engineered products and processes in the economy, from the level of material selection, which will be watched closely by NGOs in some cases, to understanding the implications of services that may be platformed on the particular technology. Of course, the sustainable engineer should not expect to internalize all this information; that would be unrealistic. But certainly one should develop the ability to ask the right questions and reach through the firm or organization to obtain the advice and information essential for the creation of a responsible, ethical, and commercially successful technology in today's challenging environment.

6.4 THE THREE LEVELS OF TECHNOLOGY SYSTEMS

Thinking about technology and technology systems is complicated in part because the relevant literature is often confused and ideological. GMO corn, for example, is attacked by activists because it allegedly supports the substitution of "industrial farming" for subsistence farming in Africa and Asia, and extravagant claims are made that nanotechnology will lead to a new era of world peace and the evolution of "posthuman" entities. It would seem that there has to be a more productive way to think about technology.

One way is to consider technologies and the goals for which they are being deployed as a single package and see what flows from that approach.[2] Take, for example, an airplane, a highly complex piece of technology. If the airplane is viewed as a means to take off in one place and land in another, it is a highly efficient and remarkably safe piece of technology. The combination of the airplane and the goal, safe point-to-point flight, can be called Level I. But how many times do you get on an airplane and it doesn't arrive anywhere near on time, throwing your schedule off? This is the airplane as part of a complicated air transport system, the goal of which is not point to point, but

[2]This model has been worked out by my colleague Dan Sarewitz and myself, and is described in more detail in *The Techno-Human Condition,* by Allenby and Sarewitz, published in 2011 by MIT Press (Cambridge, MA).

continued operations that transfer freight and people around the country and around the world relatively safely and efficiently. And in this system, with this goal, there are all sorts of things—weather problems affecting a hub airport, substitution of equipment, cumbersome and inefficient security processes, rules governing how long crews can legally be on duty—that can throw off simple point to point operations. In fact, this is a network, not a single route, and the combination of the technology and this goal create a Level II technology, one that often has unanticipated consequences, such as unpredictable delays (because, of course, that which causes them, like weather or equipment failure, is unpredictable).

But this networked transportation system, complex as it is, does not capture the Earth systems aspects of powerful technology systems. Thus, for example, the airplane technology may enable point-to-point rapid mobility (Level I technology), and global air transport networks (Level II technology), but it also has significant other dimensions. It becomes an important part of the quality of life of, for instance, aging seniors in developed countries, who have the financial resources and the leisure time to enjoy traveling (don't forget that quality of life is an aspect of sustainability, and thus of sustainable engineering). By enabling ecotourism, it becomes an enabling technology in the degradation of heretofore unreachable ecosystems (and cultural oases as well), leading to development and tourism where none would otherwise exist. By creating a network of mobility, airplane technology also knits together a global mass of human protoplasm, at least from the perspective of disease organisms, which are able to ride airplanes in real time to reach new populations to infect. Airplane technology is classic "dual-use" technology as well: commercial and military needs and technologies interact constantly as the underlying platforms evolve (as we learn to build increasingly capable unmanned flying vehicles for military purposes, for example, the day when commercial cargo, and eventually human, payloads are carried by unmanned aerial vehicles comes ever closer). This is Level III technology: technology as Earth system, affecting humans, cultures, ecosystems, natural cycles, and other aspects of the world in unpredictable, tightly coupled, and complex ways. And this is perhaps the least understood and most difficult level of technology.

Figure 6.8 generalizes this discussion, and adds the policy implications to the analysis. Thus, for example, the policy response to a Level I technology is usually relatively simple: the goal and the technological solution are tightly and completely coupled, so that the technology solves the problem. Thus, Figure 6.9, using the case study of a vaccine, shows that at Level I the technology—an effective vaccine—is entirely capable of meeting the Level I goal—prevent disease in the treated individual—leading to a relatively easy policy decision. At Level II, however, things get a little more complicated. Thus, the goal for deploying vaccine technology might now be stated as "helping the country develop economically."

Policy Response / Technology Level	Goals	Policy Response
Level I	Defined; technology and goals tightly coupled	Yes or no to technology and the explicit goal
Level II	Defined but loosely coupled, with additional interfering independent variables	Explicit decision, then deal with unintended consequences
Level III	Uncertain and often conflicting; most important may be unconscious or implicit	Real time technology and policy assessment (RTTPA)

FIGURE 6.8 Technology Level/Policy Response Matrix. This figure illustrates the three levels of technology and how they are related to goals and lead to policy responses. In general, Level I technologies are causal and relatively comprehensive responses to particular challenges, and therefore have a clarity of purposes that is diffused as one moves to Level II and Level III technologies.

Policy Response / Technology Level	Goals	Policy Response
Level I	Reduce individual's risk of disease	Agreement on goal, and therefore on technology: implement technology
Level II	Increase economic growth in developing countries by reducing costs of disease	Implement technology, but that policy alone may not lead to stated goal
Level III	Improve human well-being through vaccine technology	Real time technology and policy assessment (RTTPA) – what other systems are affected by vaccination programs; how do they respond?

FIGURE 6.9 Policy Response Matrix: Vaccine Example. This figure provides an example of how the different levels of technology play out. Note that the technology, a vaccine in this case, remains the same; what changes are the goals attributed to diffusion of the technology and the ability of the technology to contribute to achievement of the goal. At Level III, the need for a real-time response that depends on immediate system state means that no stable policy position is feasible a priori.

Unlike at Level I, this goal is not tightly coupled to the technology, although a relationship may exist. The data are not clear that vaccinations in themselves are adequate or even necessary for economic development, which may depend far more heavily on things such as effective rule of law, educational attainment, and even in the field of health, a robust and effective public health system, regardless of vaccination as a technology. Thus, some sort of explicit policy may be appropriate at Level II, but because of the loose coupling between goal and technology, planning for unintended consequences as conditions evolve, and being open to the possibility that the link between technology and policy is so weak that it isn't worth supporting, is a necessity. Thus, for example, we might implement a vaccination program as a matter of policy, but if the money required for that program begins to reduce the effectiveness of public health measures in other areas, be prepared to revisit, and possibly reverse, that decision. Note that the difference between Level I and Level II may not lie in the technology so much as in the goals the technology is intended to meet, a point discussed in more detail in Box 6.1.

Finally, there is Level III. Here, for example, we might address broad social goals such as improved human well-being, where it is clear that many different factors come into play, and that the technology, vaccination, is only a small contributor. At this level, difficult questions often arise: Suppose, for example, that a vaccination program dramatically reduces childhood mortality from a particular disease, but birth rates continue at their historical level, resulting in a population bulge of young people. Is the society ready to educate, and find jobs, for its young? Can this society provide a high quality of life and opportunities for them? Will the swell of young people support, or destabilize, important cultural and economic institutions? Did the vaccine enhance the broad goal, or did it in some ways make it more difficult to achieve? Such questions cannot be answered, in part because appropriate responses depend on how the many coupled aspects of the relevant human/natural/built systems react in real time. For that reason, **real time technology and policy assessment (RTTPA)** is the only viable option.

RTTPA is a conceptual framework for policy assessment, and for sustainable engineering, that at present is not frequently used. It is designed for use at Level II and especially at Level III where the response of the complex system within which a technology is embedded is uncertain and unpredictable. Examples might include the way that CFCs, an ideal "green chemistry" product, turned out to deplete the stratospheric ozone layer; the way the railroad generated the industrial time structures we are so familiar with today; or the way airplanes contribute to the rapid spread of disease. In all these cases, we have identified the systemic effects as the technology matured, but it would have been difficult to predict when the first few artifacts appeared. Thus, RTTPA calls for several important features to a well thought out sustainable engineering program, which includes policy dimensions of technologies.

BOX 6.1 Levels I and II, Revisited: The Shop Floor

Level I technologies are sometimes called "shop floor" technologies. To some degree, this reflects their relative simplicity: It's always easier to deal with an artifact as a physical, engineered product rather than a complicated nexus of behavioral, cultural, and institutional requirements. Yet the shop floor is just as complicated, and in a way provides a nice example of the advantages of Level I technology systems over Level II systems, and the subtle differences in goals, which sometimes make all the difference.

Consider an example from the personal experience of the author that really involves shop floors: the use of toxic chemicals in a manufacturing process, say, chlorinated solvents in a cleaning bath for small metal piece parts. If one is the individual designated by the firm to manage environmental and safety risk, this will obviously be a concern, because such solvents generally have neurotoxic and other undesirable human health impacts. An apparent way to manage the risk to employees is to provide, and mandate the use of, appropriate protective equipment. Long experience does indicate that, if properly used, such personal protective equipment (PPE) can indeed reduce the hazard of the toxic solvent by preventing body contact. But protective equipment can break down, be misplaced, or simply not used. It is not just that managers may not be as strict in requiring it as they should be because it may impede productivity (or, more frequently, they're just too busy doing other things associated directly with production to pay sufficient attention all the time). It is also that employees frequently dislike protective equipment because it is uncomfortable and makes their job harder, as anyone who has worn heavy PPE in a hot factory can attest. This is a Level II situation: the technology (PPE) speaks directly to the problem, but it only addresses it indirectly (it must be used all the time to be effective, which is not implicit in the PPE technology, but requires constant institutional and individual attention).

The Level I solution to this challenge is simply to redesign the manufacturing process to use less toxic materials in the first place. The technology—a new, nontoxic cleaning bath—directly and completely addresses the problem—potential exposure to a toxic material—by eliminating the undesirable material. Note, however, that this is not a costless approach: Changing materials such as the cleaning bath may change other design requirements in the manufacturing process, which means that a systemic balance of costs and risks may be necessary. At the least, it will require pilot testing to ensure the new process performs as specified. For example, when chlorinated solvent cleaning systems were replaced by water-based cleaning systems (one of the substitutes identified in electronics manufacture), the energy consumption of the process increased because it takes more energy to dry water from a component than to dry the equivalent chlorinated solvent. In this case, the added energy consumption (and associated carbon emissions) were considered a reasonable tradeoff for the gains in worker safety, especially because this substitution was part of a general redesign of cleaning systems then being undertaken to reduce the use of chlorofluorocarbons (CFCs) in manufacturing as well.

This approach has been formalized in the well-known occupational health and safety heuristic known as the **hierarchy of control**, which establishes a chain of potential response mechanisms, ranging from the most efficient to the least efficient, to be used in protecting worker safety:

1. Elimination of the risk (e.g., move to nontoxic water-based cleaning systems)
2. Substitution of less risky alternatives (e.g., select the least risky water-based system that still meets other operational, cost, and management requirements, even if it is still somewhat toxic)
3. Isolation of the risk to reduce potential exposure or impact (e.g., limit the number of workers who have access to the chlorinated solvent cleaning baths and materials)
4. Implementation of engineering controls (e.g., have a lock-out switch so that workers can't expose themselves by reaching into the machine, and redesign the machine bath area to be closed instead of open top)
5. Implementation of administrative controls (e.g., make part of the line manager's pay dependent on reducing worker exposure to chlorinated solvents)
6. Use of personal protective equipment

In many ways, this segues from Level I at the first-response mechanism to Level II at the sixth, as less and less of the goal (no exposure) is incorporated into the physical technology, and more and more of it comes to depend on individual, institutional, and cultural actions. As any line manager or environmental professional can tell you, the more you can reduce the chance for human error, the safer the system will be, all else equal.

From a sustainable engineering perspective, therefore, one lesson is: If the goal is relatively simple, try to design as much as possible into the technology itself; in other words, strive for Level I solutions where they are possible and effective. Where they are not, as in the example of vaccines in public health systems given in the text, try to develop the most effective combination of technology and institutional, personal, and cultural system states that will accomplish the desired goals.

First is adaptability and flexibility: Many firms reduced their use of CFCs relatively rapidly because they had technology options, and their engineers were adept enough to redesign manufacturing processes and product designs so that the non-CFC options could be implemented. Blanket bans on useful families of materials, and simplistic policy structures ("just focus on carbon footprints"), inhibit sustainable engineering and RTTPA. In addition, it is important that both sustainable engineering and related policy structures adopt a posture of constant questioning and continued engagement with the systems being modified. If a few good scientists hadn't wondered about the implications of the chemical stability of CFCs in the atmosphere, for example, we might not have caught the CFC/stratospheric ozone depletion issue in time

to respond effectively. These scientists remained engaged with the chemistry of CFCs and the systems that might be affected, even after CFCs were developed and commercialized.

If you were asked to engineer the Everglades, you might (as has been done) decide to try to reconstruct large areas of marsh—but that is only the beginning of the task. You then have to monitor the results to determine if the policy is effective: Is it, for example, reducing the flow of unwanted nutrients into the Everglades proper? More fundamentally, is it increasing or decreasing the ability of the existing flora and fauna to continue living in the Everglades? Such questions cannot be answered without continuing engagement with the system. RTTPA is a difficult challenge because many people like certainty, even when it is contraindicated by the complexity of the issues being addressed. For engineers, RTTPA is somewhat countercultural: It means that sustainable engineering is not thought of as a series of project that, once completed, are put out of mind, but instead that sustainable engineering is a series of dialogs with the technologies and the systems affected by those technologies that is more of a process than a project.

One thing that this schematic makes clear is that many people, in discussions of technology, deliberately or inadvertently confuse these levels in making their arguments. For example, many discussions of birth control technology conflate Level I arguments with Level II arguments: That condoms are effective in reducing the spread of HIV is a Level I discussion, but whether condoms are ethical in light of Christian teaching is a much more complex and difficult Level II or even Level III discussion. Even if a person believes condoms or their distribution is immoral, it does not mean they aren't effective at the limited goal of controlling HIV and reducing disease and mortality. It simply means that, to that individual, the Level I effectiveness is outweighed by the Level II or III implications. The sustainable engineer, therefore, should take care to understand what types of discussions are being made about technology, and to make sure that data and discussions of engineering choices reflect the appropriate relationship among technology, goal, and policy response.

6.5 CONCLUDING OBSERVATIONS

The discussion of technology in these two chapters has briefly covered a critical area of human endeavor and of sustainable engineering. Nonetheless, a few points stand out even in this general discussion. Most important, technological change cannot be understood as an isolated technocratic event. Especially with the disruptive technology systems that form the core of Kondratiev waves, technology represents movement towards new, locally stable, Earth systems states. These states integrate natural, environmental, cultural, theological,

institutional, financial, managerial, technological, built, and human dimensions; they even construct our sense of time and space, and they do so unpredictably and occasionally with high cost. (Note that, because social, cultural, institutional, and psychological states are usually tuned to existing conditions, any disruptive technology change will introduce major changes in those states that inevitably incur costs.) Technologies do not define these integrated Earth system states, except by convenience, but technological evolution is certainly a primary force leading to new earth system states. Again, this isn't magic: It simply means that the world with major railroad infrastructure is vastly different than the world before that technology system existed.

As a corollary, changes in technological systems can and often do destabilize existing institutions and power relationships, creating conditions leading to the evolution of new ones. Accordingly, technological change will be opposed by many who see their place in the world under siege and who quite rationally see it in their interest, if not that of society as a whole, to resist. Those who controlled the canals in the United States fought the expansion of railroads (by, for example, seeking to ban railroad bridges over the Mississippi); more fundamentally, some societies, such as China in the 1430s, chose stability and the possibility of autarky (national self-sufficiency and minimal international trade), thus leaving the world open to eventual domination by the more expansionist Eurocentric Enlightenment culture. Similarly, Japan rejected gunpowder as destabilizing of its feudal social structure, and was partially an autarky during the Edo period before it was forced open by Western trading initiatives in the 1850s.

This suggests two lessons for the sustainable engineer: First, any technology powerful enough and new enough to be interesting is likely to be opposed by those controlling the technologies to be displaced. Second, even if technological evolution can be stopped locally, it is difficult to stop any technology that provides economic or cultural advantages, especially in a globalized market. Thus, for example, the European Union despite determined efforts has been unable to stop the development of agricultural GMOs, and the United States under the Bush administration was unable to stop stem cell research.

An additional important insight is that projecting the effects of technology systems before they are actually adopted is not just hard but, given the complexity of the systems, probably impossible. Virtually nothing of what occurs at the frontiers of rapidly evolving technological systems is planned in advance, especially as these systems interact with each other and underlying social and cultural patterns. Thus, for example, no one planned the time structure that we moderns take for granted; it was an unpredictable and deep change to the time structure of pre-railroad agrarian societies, an unplanned product of our technology. Although in retrospect it is a rational, even necessary, response to the demands created by a growing and increasingly interconnected network of railroads, it

was also not foreseeable. Similarly, that automobile technology would introduce mass production, mass consumption, and mass credit was not obvious until that wave broke: The evidence is that many companies, driven by that most potent of human instincts, financial self-interest, failed completely to see it coming. They then failed economically. Technological change is a harsh taskmaster (which is why the great Austrian economist Joseph Schumpeter famously called capitalism, with its continuing innovation, a "gale of creative destruction.").

At the heart of much of the challenge of technology is the simple fact that technology is a major means by which the human will to power is reified, expressed, and channeled. Cultures that develop technology, and create frameworks within which it can react upon itself and so accelerate its own evolution, thereby gain power over competitors. Thus, for example, a recent analysis indicates that the variation in the rate and depth of adoption of technology between nations—which depends on a complex mixture of culture, attitude toward innovators, institutional stability and support, financial systems, legal and regulatory systems, educational capabilities and much else—accounts for at least a quarter of per capita income differences among countries. This is a huge difference given the compounding effect of per capita income differences over time. Because technologies create such powerful comparative advantages between cultures, those cultures that attempt to block technology will, all things equal, eventually be dominated by those that embrace it. Thus, it is likely that technological evolution will be difficult, if not impossible, to stop, despite the determined efforts of some in the environmental and sustainability communities. Indeed, whether and how technology systems can be moderated in the age of global elites becomes an important research question extending far beyond mere questions of regulation. It is partially because our time raises such challenges bluntly and unavoidably that sustainable engineering becomes so important as a framing for engineering and technology—and that sustainable engineers need to be prepared to become not just technically competent, but wise and able leaders.

Exercises

6.1 You are the chief technical officer of a software firm that plans to develop a software game program that could be platformed either on personal computers or mobile phones (but not both at this point). Compare Figures 6.2 and 6.3 regarding the market trajectories of these technologies, and determine which option they might suggest is preferable. What additional market data would you like to have before you make your decision? Does it matter where your core market is, and if so, why?

6.2 Construct a technology/policy response matrix for (a) a second-generation cognitive enhancement pill that allows full concentration

for twenty hours out of every twenty-four for use in combat zones; and (b) that same pill used in an academic setting. What do the results tell you?

6.3 You are on the ship design team that created the caravel, the ship that made European transoceanic voyaging, and thus European imperialism, possible in the fifteenth and following centuries. Following Figure 6.1, what questions would you have asked during the product realization process at each gate?

6.4 You are a member of the design team for the caravel, but have been brought back four centuries later and charged in a civil case (not a criminal case) with responsibility for the European imperialism that followed. What would you argue in your defense? Does the Chinese action of burning their fleet in 1433 help your case, and if so, how?

6.5 Can you imagine how organic agriculture might become a disruptive technology? What sorts of farming practices would it need to displace? What changes is it likely to induce in social, cultural, institutional, and psychological states?

Annotated Bibliography

1. For an overview of many of the technological issues involved in product design, see Thomas Graedel and Brad Allenby's book, *Industrial Ecology and Sustainable Engineering*, published in 2009 by Prentice Hall. The book as a whole, however, retains much of the environmental flavor that characterizes the sustainability discourse, particularly in the United States, and is therefore a somewhat interesting transition document as well. An excellent treatment of technology in the context of environmental and global systems is Arnulf Grubler's *Technology and Global Change,* published in 1998 by Cambridge University Press. Those who are interested in the history of technology might want to start with a classic text, *A History of Mechanical Inventions,* by Abbott Usher, published in 1954 by Harvard University Press; those interested in chemical engineering and the materials sector might want to start with F. Aftalion's 1991 text, *A History of the International Chemical Industry,* translated by O. T. Benfey and published by the University of Pennsylvania Press.
2. The original discussion of Levels I, II, and III may be found in Brad Allenby and Dan Sarewitz's book, *The Techno-Human Condition,* published in 2011 by MIT Press (Cambridge, MA). For a rewarding discussion of technology as social practice, see W. Bijker, T. P. Hughes, and T. Pinch's 1997 edited volume, *The Social Construction of Technological Systems* (Cambridge: MIT Press), from which the example of the caravel is drawn. Finally, for a serious philosophic treatment of technology (complete with difficult metaphysics), you might want to check out the 1977 W. Lovitt translation of Martin Heidegger's essays on technology and Being, *The Question Concerning Technology and Other Essays* (New York: Harper

Torchbooks). But beware, because German metaphysics in general, and Heidegger in particular, are not for the faint-hearted.

3. To get a flavor for the profound change implied by the scientific revolution and the ascent of technology, try reading *Faust,* the famous play by Johann Wolfgang von Goethe, together with a good history of the Industrial Revolution, such as D. Landes's 2003 *The Unbound Prometheus: Technological Change and Industrial Development in Western Europe from 1750 to the Present* (2nd ed.; Cambridge: Cambridge University Press). From the former you get a flavor of what the Enlightenment transition from human as acted upon, to human as actor, felt like; from the latter, an idea of how that dramatic cultural shift translated into a change in the human condition that laid the groundwork for the anthropogenic Earth.

CHAPTER

7 Industrial Ecology

7.1 INTRODUCTION

That the engineering profession is considering sustainable engineering as a new conceptual area of study is far from unique historically; in fact, most engineering fields that we now accept without question arose the same way. Thus, for example, civil engineering arose in the 1700s as a variant of military engineering (and was called "civil" to differentiate it from military). Environmental engineering is a recent reformulation of what used to be called "sanitary engineering"; it differs from the latter, which tended to focus on sewage and water treatment systems, in its far more inclusive scope, including such areas as hazardous material and waste management, pollution prevention, air and water contamination, and the like. Similarly, sustainable engineering can be understood as yet another shift in engineering that reflects changes in the issues and conditions affecting the broader world. In this chapter, we will explore some of the fields that sustainable engineering has grown from, and some of the tools they offer that, although not as complete and systemic as will be required, certainly provide a basis for the continued development of sustainable engineering. In particular, we will consider **industrial ecology** (and to some extent systems engineering) in this chapter; in following chapters we will address **adaptive management**, and earth systems engineering and management, both of which extend beyond industrial ecology in important ways. In all these cases, it is important to understand both the methodological strengths and potential they offer, as well as the weaknesses and caveats. Industrial ecology and adaptive management, for example, tend to focus disproportionately

on environmental and ecological perspectives, and lack sophistication regarding emerging technologies, while earth systems engineering and management offer a theoretical framework for sustainable engineering, but inadequate methodological structure at this point.

7.2 INDUSTRIAL ECOLOGY

Industrial ecology is one of a number of new areas of study, such as "green chemistry" or "ecological economics," that have arisen in response to increased social engagement with environmental issues and have deliberately tried to integrate across very different disciplines. This raises a number of difficult questions, including basic ones such as the boundaries of the field itself, for the sustainable engineer seeking to use industrial ecology principles and methodologies. Thus, like "sustainability" or even "sustainable engineering," "industrial ecology" has a number of definitions. One of the most cited ones is that provided in the leading textbook on industrial ecology:

> Industrial ecology is the means by which humanity can deliberately and rationally approach and maintain sustainability, given continued economic, cultural, and technological evolution. The concept requires that an industrial system be viewed not in isolation from its surrounding systems, but in concert with them. It is a systems view in which one seeks to optimize the total materials cycle from virgin material, to finished material, to component, to product, to obsolete product, and to ultimate disposal. Factors to be optimized include resources, energy, and capital.[1]

For purposes of sustainable engineering, there are several interesting aspects of this definition. First, because it emphasizes "sustainability" as a core of industrial ecology, this definition to some extent incorporates the ambiguity of that concept, as well as suggesting that the field begins with a normative bias (because "sustainability" is a normative concept). Second, the definition clearly requires a systemic approach, which for the engineer tends to imply that systems engineering is a potential source of methods and approaches for sustainable engineering. Third, it focuses on industrial systems, making it quite different from other environmental fields, which tend to emphasize ecosystems and "natural" systems rather than human systems, much less industrial systems. Finally, the factors to be optimized reflect an environmental bias: Resources, energy, and capital are held out, but information, which is particularly critical in areas of emerging technologies, is not mentioned.

[1]T. Graedel and B. R. Allenby, *Industrial Ecology and Sustainable Engineering* (Upper Saddle River, NJ: Prentice Hall, 2010), p. 32.

Compare the definition that was presented in the 1995 "White Paper on Sustainable Development and Industrial Ecology" issued by the Institute of Electrical and Electronic Engineers, Inc.:

> ***Industrial ecology*** is the objective, multidisciplinary study of industrial and economic systems and their linkages with fundamental natural systems. It incorporates, among other things, research involving energy supply and use, new materials, new technologies and technological systems, basic sciences, economics, law, management, and social sciences. Although still in the development stage, it provides the theoretical scientific basis upon which understanding, and reasoned improvement, of current practices can be based. . . . It is important to emphasize that industrial ecology is an objective field of study based on existing scientific and technological disciplines, not a form of industrial policy or planning system.

This version of the definition explicitly emphasizes industrial ecology as an objective, not normative, pursuit. Moreover, as we would expect from an engineering professional society, it mentions new technologies and technology systems as important factors to be incorporated in industrial ecology. But it leaves hanging an important question, which the industrial ecology community itself still struggles to answer: Is industrial ecology intended to support the social goals of sustainability, such as egalitarianism, or is it intended to be an "objective field of study"?

The essence of industrial ecology, and the source of the analogy itself, is that ecosystems evolve to generate more efficient use of resources. From the first primitive soup where a single type of organism simply eats what it finds and excretes waste back into the environment (a Type I system in Figure 7.1), there is movement to a system where the waste from one organism is used by another as food before waste is produced (a Type II system), then finally to an ecological community where resources are cycled within the community before eventually becoming waste. As indicated by the size of the spheres, each jump in function is accompanied by the ability to support a greater degree of biomass with the same input of resources. This is obviously a highly oversimplified and schematic version of both biology and economic activity, but it has the virtue of clarity and prescriptive simplicity: If you want more economic production, increase the cycling of materials within the economy before you finally dispose of a material or product. Perhaps the biggest issue hidden by this simple model is energy demand. Each jump in complexity requires more energy, all things equal. Thus, if energy is the limiting factor of an economy, either because we can't get enough of it or because the necessary sinks are inadequate (e.g., climate change and fossil fuels), there is a limit to the internal cycling that can be done.

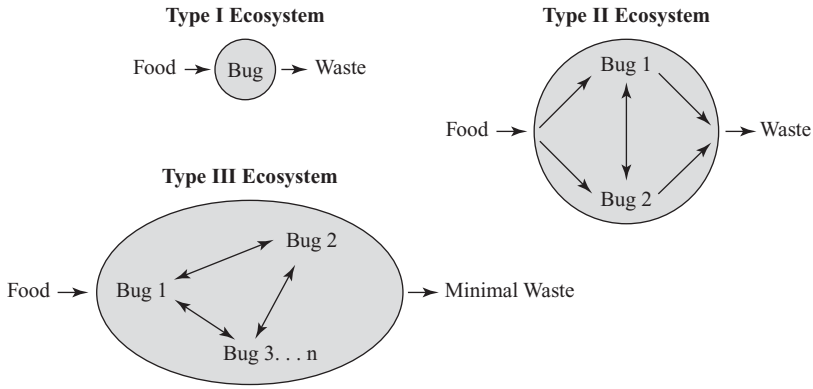


FIGURE 7.1 Industrial Ecology Conceptual Model. This is a simple schematic presenting the idea behind industrial ecology: If one designs an industrial or economic system to mimic elements of an ecological system, one can obtain greater function with less marginal investment of material and energy. In this case, each jump in internal cycling of materials results in the ability to support a larger biomass, all else equal. Industrial ecology seeks to apply this basic insight to the economy, at least as an overall organizing principle.

Obviously, this first insight is just the beginning. Industrial ecology as it has evolved reflects several fundamental themes, some of which resonate with the sustainable engineering themes we have identified in this text. Among the most important are:

1. Industrial ecology emphasizes the system, rather than specific engineered or natural elements of the system.
2. Industrial ecology tends to emphasize long-term concerns that are of regional and global scope, and persistent and difficult to manage. This emphasis probably arises because industrial ecology is trying to establish a boundary between it and existing fields of study such as environmental engineering.
3. Industrial ecology, although highly multidisciplinary, tends to emphasize scientific and engineering analysis, and focus on environmental issues.
4. Industrial ecology attempts to understand and protect the resiliency of natural and human systems. This is easier said than done, because in some ways the property of resilience is as complicated as the concept of sustainability.
5. Industrial ecology uses such systems techniques as mass-flow analysis, DFE (Design for Environment), and LCA (Life Cycle Assessment) to understand energy and material flows through economic and environmental systems, and the linkages among them.
6. Industrial ecology views private firms and industrial systems at multiple scales as central to managing environmental impact. In this, industrial

ecology differs from traditional environmental fields, but it also betrays an abiding bias towards environmental effects rather than the broader issues with which sustainable engineering engages.

7. Industrial ecology studies are done at many different scales, from looking at a particular material or product, to a facility scale, to broader regional and global scales. Each level of the system has its own issues, and the appropriate level for analysis of a particular issue varies. Thus, for example, Figure 7.2 illustrates several of the relevant scales for automotive technology. If one is working on making the engine more efficient—say, by introducing a supercharger that provides equivalent power with less fuel consumption—one need not consider the implications for the built environment. But if one is introducing an electric plug-in vehicle, one needs to be seriously concerned about the state of the infrastructure, in particular whether the grid has adequate capability to meet increased demand for power.

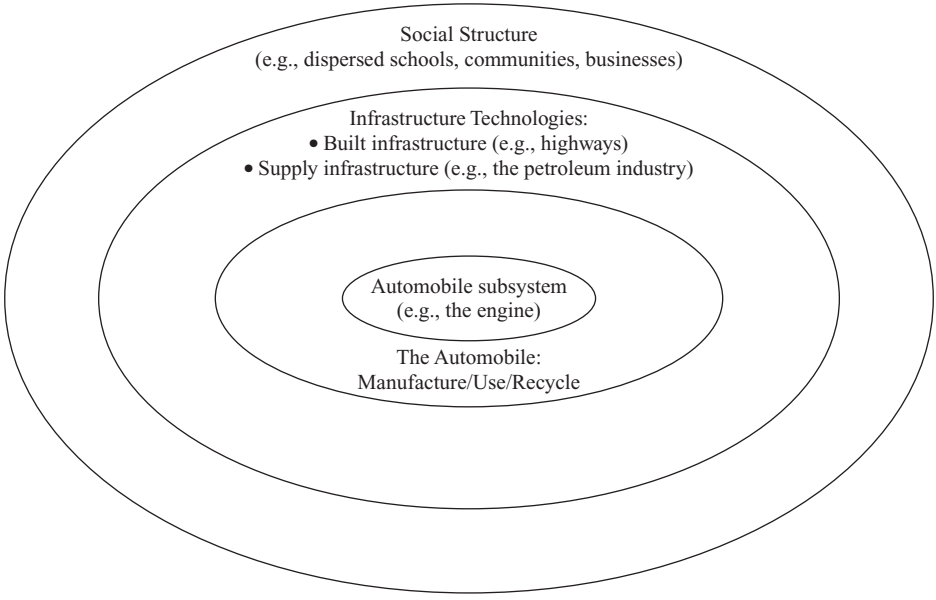


FIGURE 7.2 Industrial Ecology Systems Hierarchies: Automotive Technology System. This schematic presents an industrial ecology perspective on automotive technology. Which level is appropriate for a particular analysis is determined by the nature of that analysis. In addition, a potential improvement at one level—say, substituting aluminum for steel in a car body—might have effects at another level—reducing the economic incentive to recycle car bodies, which is dependent on the amount of steel recovered—that should be part of a sustainable engineering design analysis.

(Based on: Graedel, T. E. and B. R. Allenby. *Industrial Ecology.* 2nd ed. Upper Saddle River, NJ: Prentice Hall, © 2003, p. 69. Reprinted by permission of Pearson Education, Inc.)

In many ways, these themes, and the way industrial ecology has been developed, are in large part a reflection of its history. Students are often given the idea that internal consistency and applied rationality govern the creation of knowledge, but the field of industrial ecology is young enough so that the interplay of social, economic, and even personal biases and pressures can still be seen. Consider, for example, the history of the term itself. *Industrial ecology* seems to have been first used in the title of a very limited circulation journal in 1970, only one issue of which apparently was ever printed. It was most probably unsuccessful because at that time, the relevant publics—environmental activists and NGOs, academics in environmental disciplines, environmental professionals in firms—were not yet ready to consider environmental issues as important for industry to consider in their strategic activities, such as product and process design and management. (Waste streams, of course, were managed to some extent, although subsequent experience showed that they were not being managed well.) Although it is hard to believe now, even the coupling of "industrial" and "environmental" struck many people as so oxymoronic as to be not worth even contemplating.

But the biological analogy implicit in the phrase "industrial ecology" proved very tempting. In Japan, the Ministry of Industry and International Trade in 1972 began considering such an analogy as a model for structuring the Japanese industrial system, although this program appeared to fade with the 1973 energy crisis. By the 1990s, however, a series of articles and books had taken up the theme, and this time society and the relevant professional communities had developed their thinking to the point where the term became fascinating rather than laughable. This third, and ultimately successful, attempt to introduce industrial ecology began with a 1989 article by Frosch and Gallapoulos, which suggested that industrial systems could be more efficient if their material flows were modeled after natural ecosystems. This suggestion was interesting in part because both authors worked for a large firm, General Motors, that had not been notably associated with environmentalism up to that time. Textbooks, a journal, and a society, all important steps in institutionalizing a field of study, soon followed.

It is not just the methods of industrial ecology, then, that the sustainable engineer should understand, but the process by which industrial ecology succeeded. What conditions made it possible—and are similar conditions in place for sustainable engineering? For one lesson of industrial ecology, drawn from its history rather than its disciplinary knowledge, is that even a good idea will not blossom if the culture, institutions, and worldviews are not receptive at the time.

In this regard, it appears clear that industrial ecology arose when it did at least in part as a reflection of a deepening understanding of environmental issues. In particular, the complex, systems-based nature of regional and global

environmental perturbations such as ozone depletion, regional and global spread of pollutants, and loss of biodiversity, and the inability of reductionist regulatory approaches to address such issues adequately, became increasingly obvious. As industry saw it, rising regulatory costs and responsibilities and the demands of stakeholders and customers encouraged new, more systemic ways of thinking about these issues, partly to avoid unnecessary and steeply increasing costs, and in part to avoid unanticipated disruption of operations. This latter issue was highlighted for industry when the movement to ban chlorofluorocarbon (CFC) use accelerated, because unlike end-of-pipe regulations that could be implanted without affecting manufacturing operations, CFCs were used in many core manufacturing and product design activities. It was thus a direct and unexpected challenge to many existing processes and products, requiring in some cases serious redesign of existing systems to continue operations.

More operationally, industry, particularly the electronics industry in the United States, began to develop a set of practices based on industrial ecology principles, generally grouped into a set of "Design for Environment" methodologies under their Design for X (DFX) product realization processes (see Chapter 6). The initiative in product-oriented approaches subsequently shifted to Europe, where a focus on product takeback and significant chemical regulation encouraged the development of underlying industrial ecology tools, and which today is a leader in consumption related regulation. The Netherlands in particular took a broader approach that began to move away from a purely environmental focus, seeking to define a sustainable society in one generation; the country published its first *National Environmental Policy Plan* in 1989 and the *National Environmental Policy Plan Plus* in 1990.

This history explains several aspects of industrial ecology as currently practiced. First, reflecting its origins, industrial ecology and many of its methodologies, such as DFE and LCA, remain focused to a large extent on environmental issues, although there are continuing attempts to broaden its research agenda and scope. This is in large part because many of the people involved in the field came from environmental backgrounds, which is overwhelmingly where the activism and regulation were focused.

Second, much of the early industrial ecology literature tended to focus on manufacturing sectors. Again, this is because such firms were the ones most affected by environmental regulations and politics at that point, and because much environmental impact was seen as coming from manufacturing. Interestingly enough, that industrial ecology was developed at first primarily by firms initially created resistance among environmentalists, some of whom viewed industrial ecology as an effort to co-opt activist environmentalism. Although this concern may have lessened over time, it still remains true that a more subtle effect of this history might linger in the strong continued focus on

manufacturing technology and engineering approaches in industrial ecology. More subtly, because this approach combines an emphasis on traditional manufacturing with environmental interests, it tends to embed a focus on material and energy flow, making industrial ecology fairly poor at dealing with questions of emerging technology, not to mention the service sectors of developed and developing economies.

Third, industrial ecology was developed in advanced economies to respond to environmental problems as perceived and valued in such countries; it thus has not been particularly effective or even useful in developing economies. Moreover, it is mildly ironic that Europe can probably be considered the leader in industrial ecology at this point, when manufacturing has so strongly shifted away from Europe and the United States towards Asia.

Finally, and perhaps most subtly, industrial ecology tends to accept the analogy between biological systems and human systems implied in its title far too readily: a modern corporation is far more complex than a natural ecosystem in important ways (e.g., **wicked complexity**); unless care is taken, the overuse of biological or ecological metaphor can be as inappropriate as any other analytical deficiency.

These are all cautionary lessons for sustainable engineering. Most obviously, there is a tendency for sustainable engineering to slip back into being just "green engineering," which is inadequate. Second, it is far easier to think about and design purely physical systems, and worry about materials and energy, rather than to try to manage complicated human/built/natural systems, such as military technologies in the real world (Chapter 11), or the Everglades or the Aral Sea (see Chapter 13), or the services that increasingly dominate economic activity. The panic of 2008, for example, raised a host of challenging sustainability questions and involved many engineers (financial firms like to hire engineers because they know them to be quantitatively competent), but it was not on the radar of any industrial ecology or sustainability study. Sustainable engineering, and the requirements of the modern world, do not allow for the luxury of ignoring the world as it really is. Finally, sustainable engineering requires problem solving in particular contexts: What is sustainable engineering in Cambridge, Massachusetts, or Berkeley, California, is not sustainable engineering in a village in Ghana or Mozambique, and visa versa. It is a world of significant contrasts and the state of infrastructure, of education, of technical skill, of social and cultural context, is always important to the engineer. Knowing which tools to use where, and being sensitive to which engineering skills may be most useful for a particular project in a particular locale, or for a particular product, is an important element of sustainable engineering.

As part of the development of the field, industrial ecology methodologies have blossomed over the past few years, from initial efforts at "green design" to LCA and DFE. Many of the techniques and methods, however, tend to have

certain common features that, taken together, constitute a sort of industry ecology model. Most important, perhaps, is the systems orientation. This does not mean that every analysis should include all possible impacts on any relevant system, which is often impracticable. But it does mean that boundaries drawn around particular systems or inherent in methodological approaches should be justifiable and reasonable in light of the purpose of the analysis and the nature of the system itself. Thus, for example, if a sustainable engineer is designing a new airplane and wants to make a reasonable choice between composites and aluminum for the skin, evaluation of the role of air travel in spreading disease organisms is not necessary. But it may well be necessary to consider the effects of the material choice on the supply chain, especially if a particular choice would substantially increase demand for, as an example, mining activities in developing countries.

Like sustainable engineering, industrial ecology is a young field, so it is instructive to look at some of the definitional questions that have arisen. One that the industrial ecology community has struggled with for a while is perhaps the most basic: trying to specify the boundaries of the field of industrial ecology, especially the extent to which various social sciences are included within the field and the extent to which social science research and approaches should be integrated into industrial ecology studies. In other words, if we drew a line around "industrial ecology," would social sciences be inside or outside? Similarly, in part because it has been derived from the environmental movement, there is a question as to whether industrial ecology is a field of study whose purpose is to support sustainability and environmental activism, or whether it is supposed to be objective and scientific. Is its purpose to support activism, or to provide objective data that informs but is not intended to take the place of policy debate? These questions also arise with sustainable engineering, although because it is an engineering domain, and thus more tied to problem solving in the real world, the opportunity for activism is lessened. It is by no means absent, however; many sustainability advocates seem to believe that sustainable engineering is only valid to the extent it supports their activism.

Another interesting question is whether the name of the field, "industrial ecology," should be taken as an analogy or as a mandate. Most industrial ecologists seem to understand themselves as using biological systems, often characterized by complex flows and transformations of energy and materials, as a useful source of insight and potential models—the analogy approach—but some seem to go beyond this and perceive a design mandate. For sustainable engineers, this is probably a caution: We should seek inspiration wherever we can find it, but we should avoid getting locked in to a particular worldview or ideology, especially where that is assigned a priori probative value. That a particular approach or model comes from biology or ecology might make it useful or instructive; it does not make it somehow magically better.

7.3 LIFE CYCLE ASSESSMENT

LCA is a major industrial ecology methodology, and in some ways exemplifies both the strengths and weaknesses of industrial ecology. As a tool, it reflects perhaps the most important conceptual jump made as industry internalized environmental issues: the idea that simply thinking about a particular lifecycle stage of a material or product was no longer good enough, and the entire process needed to be considered. (See Figure 7.3 for a schematic of a typical LCA lifecycle.) In practice, this idea helped shake environmental professionals out of a single-minded focus on waste streams, which was reinforced by the early emphasis on air and water pollution, and into a more systemic exploration of the relationship between economic production and natural systems.

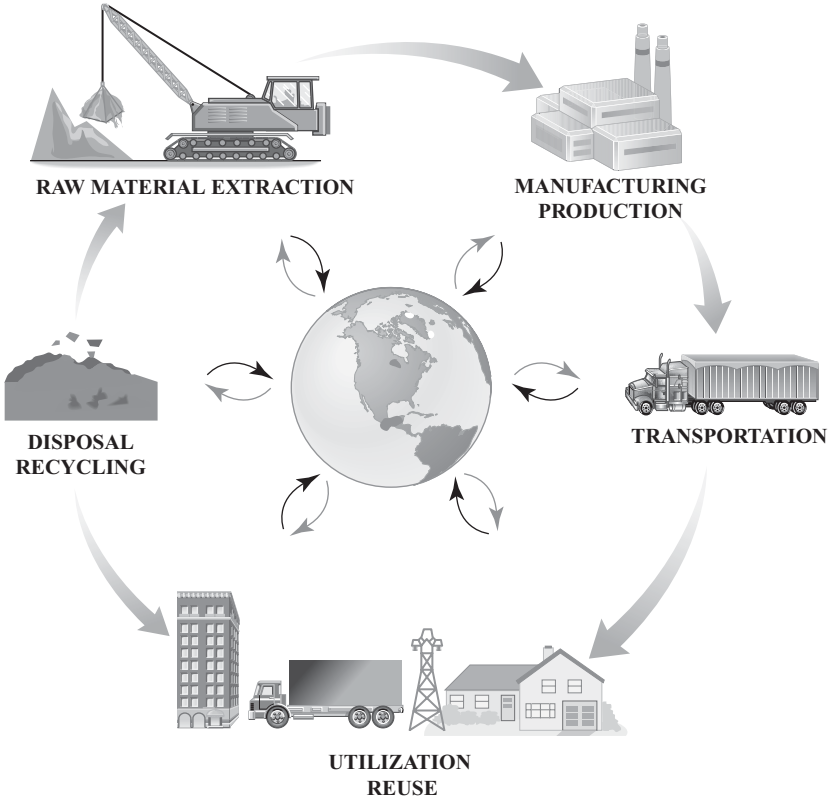


FIGURE 7.3 Simplified Lifecycle of a Material or Product. LCA would attempt to identify, and quantify the environmental impacts at each stage, thus providing data enabling systematic and rational improvement of environmental performance.

(Source: National Institute of Standards and Technology.)

But because of the power of this idea, it has also tended to inhibit industrial ecology theory and practitioners from moving further to consider technology as a complex system, and at least integrate services and social systems and values as well. Thus, although the sustainable engineer should appreciate LCA as a tool and as a concept, LCAs are not by themselves sustainable engineering.

To begin with, LCAs are at heart a mechanism to identify and quantify the material and energy flows involved in the lifecycle of a particular product, process, or material, and to identify and quantify the resultant environmental impacts. These data are usually summarized in a relatively small number of metrics or a single figure of merit, and then used to evaluate alternative approaches that reduce or mitigate environmental impact. The studies have grown quite complex, and data acquisition can be a problem, but the use of LCA methodologies of various kinds, many available on the Internet, has grown rapidly.

Whether use of LCA has actually resulted in meaningful environmental benefit in regional or global systemic terms is difficult to say, because the systems involved are so complex. A product may be designed for lower energy consumption, for example, but the more complicated design may increase demand for a material mined in a country governed by a corrupt kleptocracy, which uses the resulting increase in funding to build huge and wasteful vanity infrastructure (palaces, for instance) with a substantial environmental impact—or to build an army of thugs that further brutalize the population. That does not, however, mean that LCAs aren't useful. For one thing, they can identify areas where improvements can be made *all else equal*, which is usually worth considering; for another thing, simply using LCAs increases sensitivity to the broader impacts of design decisions, which is in itself a useful shift in individual and institutional perspective.

A number of organizations—ranging from the Society of Environmental Toxicology and Chemistry (SETAC), which did much of the original work on LCA development, to the International Standards Organization (ISO), to consulting firms that often keep their methods and data proprietary, to academic institutions—have produced LCA methodologies and data sets. In general, a three-stage method has evolved.

The first step is to define the goals and scope of the LCA. This stage is often the most critical, because only at this stage, before actually performing the LCA, can we consider potential alternatives in some depth. For example, if we are looking at an electric vehicle as an environmental technology, alternatives might include gasoline and diesel vehicles; but if we are looking at alternatives for personal transportation, we might want to consider bicycles, mass transit options, walking, virtual offices, and other functional alternatives. In the first case, the boundary set on the problem implicitly suggests motorized personal transportation that will work on existing infrastructure (e.g., roads

and highways); in the second case, we may need to consider impacts not just on the mobility technology itself, but on infrastructure. Bicycles or walking, for example, may be preferable to even an efficient electric vehicle, but may require considerable addition to existing infrastructure, such as bike lanes and walking paths, which entail different political and cost implications.

To take another example, assume we want to reduce the environmental impacts of dry cleaning as a process. If we assume existing facilities as a given, we might only include good housekeeping techniques (which frequently can yield significant reductions in emissions, at least in the short-term), or changes in emission control technologies, or in the way materials are handled. But we can also take a functional approach. First, ask why the particular activity is being undertaken. With dry cleaning, for example, studies have indicated that many people send clothes out for cleaning because they need pressing or are wrinkled, not because they actually need to be cleaned. More fundamentally, we might note that "clean" is a relative term that varies across economically similar cultures. The French, for example, tend to be less compulsive about cleanliness than the Americans. This might suggest a number of functional alternatives (provide pressing services as an independent option, thus reducing unnecessary use of cleaning solvents, or perhaps try to reduce culturally mediated demand for cleaning). We might also note that modern alternatives for both cloth and clothing construction should be able to eliminate the need for dry cleaning through proper material selection and product design, a functionality option that has yet to be really considered seriously. And, of course, we would want to investigate less complicated alternatives such as identifying less toxic chemicals that can be used with existing fabrics and clothing designs. Thus, existing dry cleaning operations almost all use tetrachloroethylene (also known as perchloroethylene, or "perc"), a chlorinated solvent that has potential human health effects; alternatives might include supercritical carbon dioxide or water-based solvents, and even more exotic possibilities such as silicone. Each has benefits and costs; a sustainable engineering analysis of such options is complicated by the fact that the costs and benefits may accrue to different groups of people. Moreover, the costs and benefits may arise in different domains: perc may have environmental costs, but the industry knows how to use it, it is relatively cheap, and it cleans well. Many of the alternatives do not clean as well, are more expensive, and require more energy. Which is best under such circumstances is not easy to determine.

This step has two important implications for the sustainable engineer. First, pick the analysis that matches the degrees of freedom you really have. If you are redesigning a cell phone for use in a network already in place, you may be considerably constrained by the existing designs of both the cell phone you are replacing and the network itself. Although there may still be things that you can do that are worthwhile, such as designing for energy-efficient operations,

and to enable software rather than hardware upgrades, thus reducing the amount of unnecessary waste produced, there are clearly things beyond the scope of your problem. For example, the question of whether a chip set embedded in your head could replace a cell phone would not necessarily be a useful one to consider. But if, say, you are designing the first iPod, then you have far more flexibility, because you are building not just a device but a service that does not yet exist. Still, you are constrained: coming up with a product that does not match human ergonomic needs would be unacceptable, and you have all the usual constraints—cost, consumer acceptance, time to market—that limit what you can pragmatically suggest.

The second implication is that, with any LCA or similar analysis, one must pay close attention to the assumptions that go into framing the study. Many commercial LCA packages, for example, place a very heavy emphasis on energy efficiency or even carbon footprint, reflecting society's current fixation on climate change. This may be appropriate in some cases but if, for instance, you are working with a product or process where toxicity is a major concern, such methodologies may give you seriously dysfunctional results because they emphasize energy efficiency over all other values. More subtly, in some cases practitioners may skew LCAs to get the results that their clients want by adopting the appropriate assumptions; it is not that the assessment or the use of the method is "wrong," just that the complexity of the systems involved allows such results. For example, if you are studying whether a plastic or a paper cup is better, and you assume that a major factor should be biodegradability, you will favor paper; if you do the same assessment, but assume energy efficiency is better, you will favor plastic. The potential for "gaming" LCA processes is apparent; companies may want their products to appear to be preferable to those of competitors, or environmentalists may want to disparage popular products to drive down consumer demand.

Once you have decided what the scope of your LCA will be, you then perform what is known as an "inventory analysis" or "lifecycle inventory analysis." This analysis simply looks at material and energy inputs, and identifies the emissions and other outputs, at each lifecycle stage. The result is usually some sort of table implicitly or explicitly coupled to a flow diagram of the relevant lifecycle. (Given the way LCAs are usually used, this generally focuses on a manufactured product, or a material in a product, or less often, a process.) Validation of the physical aspects of an LCA usually involves mass balance assessment: The total mass of an input material, for example, should equal the material embedded in the resulting product, plus the emissions to various media.

The third stage of an LCA is the "impact analysis." Here, the inventory of emissions is coupled to the resultant environmental impacts. This is a difficult process, in part because the necessary data are often either incomplete or

frequently nonexistent. It is also controversial because it is here that different impacts are weighted differently for purposes of producing high-level scores or final metrics. An extreme example of this is the "carbon footprint" LCA, which reduces all environmental impacts to implications for global warming. Using such LCA techniques, activists have tried to justify everything from proposed multi-thousand dollar charges for having a baby (because each baby is a new carbon footprint) to demanding that people get rid of their pets (based on studies indicating a pet dog has a larger carbon footprint than an SUV). Regardless of what we may think of such claims, they do illustrate how difficult, and even misleading, aggregation of impacts of different kinds can be. (Do people really buy pets based on their carbon footprints?)

The final step, of course, is to use the LCA results. This is trickier than it sounds, because it involves not just understanding the technical strengths and weaknesses of the particular analysis that has been done, but the relevance of that analysis to a larger policy or industrial context. So, for example, in the late 1990s AT&T, which at the time was a major manufacturer of electronics products, did streamlined LCAs on a number of its products and found two areas that consistently scored fairly high as far as environmental impact: the use of lead-based solders in electronics products, and packaging of both incoming components and supplies, and outgoing finished products. Addressing the first issue turned out to be unfeasible both technologically and economically; in fact, the electronics manufacturing industry spent many years developing workable alternatives to lead-based solders. Addressing the second was possible within the existing technological and cost constraints, and actually was presented and implemented as part of an ongoing company focus on cost-containment initiatives. Many technically oriented people can do LCAs; it takes a good engineer, versed in the sorts of skills required by sustainable engineering, to use the results effectively and well in the real world.

Managing effective use of LCA techniques, then, requires that the scoping and the implementation of the results be considered part of the same activity. At the beginning, you might want to explicitly consider why the study is being done at all, how the results will be used, and by whom. Thus, for example, if a study is being used for internal improvement of a product, it can be done using streamlined techniques, the documentation can be informal, and the assumptions can be chosen based on professional judgment rather than popular whimsy. But if the study is going to be used in a public marketing campaign for a product, it has to be done in such a way that it can be defended: The assumptions should be clear and the study needs to be well documented. Moreover, in the former case the LCA can be done by technically competent professionals acting alone; in the latter case, corporate lawyers and marketing people should be involved from the beginning. If, for example, marketing tells you that the only meaningful metric being

demanded by the customer is a "carbon footprint" of some sort, then the LCA will be profoundly different than if you were seriously interested in the many environmental implications of a product.

These examples also make another point about LCAs: There are many different specific LCA methodologies. Perhaps the most common is founded on a bottom-up approach, where you goes through the three stages of an LCA based on the specific characteristics of the material or product in that particular application, developing an analysis using the data relevant to the case at hand. Because these sorts of studies can be unnecessarily detailed, another very common form of LCA is the "streamlined" version, where the analysis is done at a fairly high level, and is designed not so much to produce detailed quantitative data as to identify and qualitatively rank the most important impacts. This is frequently done where you wish to gather information on a number of products in a short period of time; it is especially popular with companies with short product turn-around times. A third category with some similarity to the streamlined LCA is the top-down; the most common form here is an LCA based on the economic input/output models that have been developed by economists for most countries. (Essentially, an input/output model is a huge matrix consisting of the major sectors in an economy, and the economic flows between them; they allow you to make projections by, for example, changing output in one sector and seeing what effect that has as it ripples across the economy.) Thus, for example, the Carnegie Mellon University Green Design Institute has over the years built up an LCA engine, the **Economic Input-Output Life Cycle Assessment (EIO-LCA)** methodology, available on the Net (www.eiolca.net), which uses such economic models, augmented by environmental data, to build a top-down, economy-wide, LCA model. Because use of different LCA techniques produces very different answers, the use of "hybrid LCAs" is becoming more widespread. Here, the results of a top-down assessment, using something like the EIO-LCA model, are combined with a bottom-up process or product LCA.

Regardless of the particular technique chosen, the sustainable engineer should approach LCA methods with appreciation for their ability to identify issues or implications that would otherwise have been missed, but due caution for their limitations. Many of these arise from the basic conundrum that LCAs are used broadly, and thus must have some uniformity of assumptions, underlying data, and valuation, but environmental impacts are local.

Perhaps the most important deficiencies involve data. Data are almost always partial and not location specific; moreover, a lot of the data underlying these methodologies is severely dated. Especially when environmental, economic, or technological domains are changing rapidly, this adds uncertainty to LCA results, no matter how quantitative and rigorous the particular technique. Similarly, where local conditions differ significantly from those built into the

LCA database, results will be less viable: If we are concerned about environmental impacts in a city with a serious smog problem, for example, the results of a typical LCA may be marginally helpful.

Likewise, virtually all LCA methodologies assume linear system behavior. This, of course, amounts to an assumption that we are dealing with simple systems. In some cases—for example, if we are assuming a relatively incremental change in the status quo—such an assumption may not be problematic. In other cases, however, we can rapidly get into nonlinear behaviors. For example, an incremental increase in corn-based ethanol production in the United States could have been modeled, and the assumption of linearity would not be a serious problem. But when public policy created a large jump in demand, the ripples across other systems, such as shifts in crop choice by farmers, had huge, and unpredictable, impacts across many other systems. In other words, that kind of change entered seriously nonlinear territory, and any attempt to model it using assumptions of linearity would have produced incorrect results.

In addition, LCAs are limited in practice to material and energy factors. Bottom-up LCAs are very poor at considering services, or even the service dimension of products and materials, while the EIO-LCA-type model does consider services (because models of economies necessarily encompass services), but does so only in the relatively unsophisticated way that economic data break down the economy between sectors. (Is a highly sophisticated integrated pest management system, based on large information flows and satellite monitoring, a service, or simply agriculture as usual? Is the impact of a cell phone network based on the energy consumption of the cell phones, or on the economic and social practices that it enables? Is a virtual office defined by the energy consumed by the computer, or the additional economic activity that virtuality enables?) Not only do they tend to deal partially with services at best, but services raise very different definitional problems. LCAs, after all, depend on being able to define a lifecycle: What is the lifecycle of the Internet? Of cell phone communication? Of virtual reality?

So what should a sustainable engineer make of LCAs? First, that taken as a whole they are a useful work in progress, but they are not the definitive quantitative solutions that many practitioners and consultants claim they are. The data deficiencies, reliance on simplistic mechanisms for integrating different impacts into higher level metrics, and the essential impossibility of imposing a simple system methodology on underlying complex systems with absolute rigor all suggest that LCAs are best regarded as scenarios, rather than as mechanisms for obtaining quantitatively rigorous answers. Thus, we are often able to identify important impacts or potential issues that might arise through use of an LCA, which does indeed make better process and product design possible. This is quite useful. Moreover, the development of

LCAs can serve as a guide for developing similar tools for sustainability analysis (albeit such tools will need to be significantly more sophisticated than LCAs conceptually).

But from the viewpoint of sustainable engineering, LCAs have two major deficiencies. First, they embed the environmentalist focus on material and energy, and are thus inadequate for truly systemic analyses, especially given emerging technologies and the dominance of services and virtualization of many important social and economic functions. Second, despite some claims to the contrary, they are methods that work best on environmental issues, and have virtually no social, cultural, or even economic dimension; they may thus be part of, but they are by themselves inadequate for, any sort of real sustainability analysis. There are some recent initiatives in Europe to create sustainability LCAs, but given the contingencies of existing environmental LCAs, and the normative complexity of sustainability, these should be regarded as interesting efforts worth understanding and exploring, but not ready for application to real-world issues.

7.4 STREAMLINED LCA

In many cases, a full LCA will be either unnecessary or impractical. In such cases, the use of various screening methodologies, designed to help identify first-order issues but not for rigorous analysis of particular options or support of comparisons between options, may be the best choice. Because these methods are usually at least partially qualitative, and generally can be done using at a higher scale, and using less time and resources, than traditional LCAs, they are sometimes classed as "streamlined LCAs."

Such methods often use matrix structures. Especially if we are interested primarily in identifying major considerations of complicated technologies or systems, where traditional LCAs may be difficult to apply or rely on dated information, such streamlined techniques can be an appropriate balance of rigor, analysis, resource expenditure, and most important, professional judgment.

There are obviously many ways in which such matrices can be constructed, depending on issues of concern, the types of technology or industrial systems at issue, and how detailed a matrix needs to be constructed. In all cases, of course, we would want to have an inclusive set of potential issues, and cover the lifecycle of the product, process, or service; otherwise such an assessment might be not only streamlined, but unsystematic and partial.

One such family of matrices, which enables integrated analyses across the multiple scales that often characterize industrial systems, is presented below. The family begins with a simple matrix focused on material selection considerations (Figure 7.4), which can be used to inform matrices at the process and product level (Figure 7.5), which form part of the information base for evaluating

Lifecycle Stage	Environmental Concern				
	Ecological/ Biological	Energy Use	Solid Residues	Liquid Residues	Gaseous Residues
Initial production processing					
Secondary production/processing					
Application: Manufacturing stage					
Application: Use stage					
Disposal, recycle					

FIGURE 7.4 Material Environmental Evaluation Matrix. This figure illustrates a matrix that might be used for high-level analysis of a particularly important material used in the manufacturing process or designed into a product. An "important" material might be one that is particularly toxic, or used in high volumes, or critical for continued production, or of high political or stakeholder concern (such as columbite-tantalite [coltan] or a product of African conflict zones).

Life Stage	Environmental Concern				
	Materials Choice	Energy Use	Solid Residues	Liquid Residues	Gaseous Residues
Preproduction					
Product manufacture					
Product packaging and transport					
Product use					
Recycling, disposal					

FIGURE 7.5 Environmentally Responsible Product and Process Mix. This figure illustrates a matrix that might be used to evaluate a product or a manufacturing process. Note that it can incorporate the results of the previous material analysis as one of the environmental concerns. For a process, one evaluates the environmental concerns associated with production of the process equipment and the materials (e.g., solvents, binders, lubricants) required to operate the process machinery. For many processes, the major impacts are in the process use and recycling/disposal lifecycle stage, especially where processes involve solvents. Recycling materials and chemicals used in manufacturing processes is a traditional, and still important, environmental concern.

Lifecycle Stage	Environmental Concern				
	Ecological Impacts	Energy Use	Solid Residues	Liquid Residues	Gaseous Residues
Site selection, development, infrastructure					
Principal business activity—products					
Principal business activity—process					
Facility operations					
Facility refurbishment, closure, transfer					

FIGURE 7.6 Environmentally Responsible Facility Assessment Matrix. This figure illustrates a matrix that can be used to evaluate industrial operations at the facility level. As in other matrices in this set, it can include the result of lower level assessments in the "Principal business activity—products" and "Principle business activity—process" categories, which in turn can include the material assessments. The last lifecycle stage is seldom thought of in current facility design, but is an important part of sustainable engineering. Designing infrastructure, including facilities, that are easily updated and freshened, or converted to other uses, is an excellent way to reduce unnecessary construction activity and, more subtly, support continued value of lands and communities within which industrial facilities are located.

facilities (Figure 7.6), which in turn can be integrated into an assessment of the relevant network of facilities, the industrial infrastructure (Figure 7.7). Figure 7.8 provides an idea of how these relative rankings can then be integrated into an overall systemic assessment.

There are a few obvious caveats about such a system. First, it provides an interesting and accessible example of how we might go about a multiscalar assessment process; this is particularly important because multiscalar implications are the rule, not the exception, as we begin to grapple with the complexity of sustainable engineering challenges. But note that it is the professional's knowledge and skills that really knit together the explicit matrix system; the method is useful, but not sufficient, for integrated analyses. Moreover, it would be a mistake to assume quantitatively valid linkages between different matrix levels; remember that such a system is better thought of as a guide to professional judgment, rather than a source of valid analysis in itself.

There is also the question of quantitative versus qualitative judgments, an important issue that many methodologies, by using highly quantitative

Lifecycle Stage	Environmental Concern				
	Ecological Impacts	Energy Use	Solid Residues	Liquid Residues	Gaseous Residues
Site selection and preparation					
Infrastructure manufacture					
Infrastructure use					
Impact on ancillary systems					
Infrastructure end-of-life					

FIGURE 7.7 Environmentally Responsible Infrastructure System Matrix. This figure illustrates a matrix that can be used to evaluate industrial infrastructures, including networks of facilities. At this point, reliance of streamlined LCAs on professional experience and judgment becomes apparent, especially in general categories such as "Impact on ancillary systems." No large infrastructure operates in a vacuum; there are always interconnections with ancillary systems that should be considered where appropriate. Mining operations, for example, create significant demand for energy, transportation, and water infrastructure. Such networks, of course, are highly situation dependent.

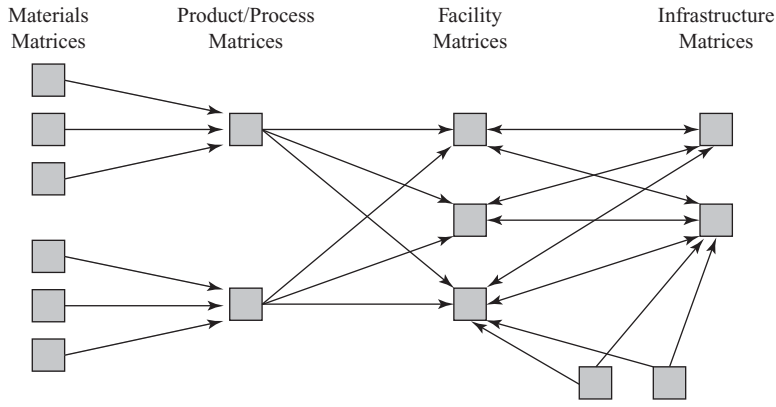


FIGURE 7.8 Matrix Evaluation System. This figure shows schematically how a network of analyses can be strung together to provide a high-level, and inherently qualitative, assessment of an industrial system. Because such techniques are heavily dependent on judgment, normative weighing of different considerations, and situation, they should not be relied on for legal or marketing purposes. They can nonetheless be very useful in gaining a quick and surprisingly structured sense of the environmental impacts of industrial systems.

mechanisms, hide but do not resolve (this is true of virtually all LCA techniques). It is true that the matrix format suggests analytical rigor; the use of 25 cells and a rating system of "0" through "4," with "4" being the most preferable, can even be used to provide a figure of merit (e.g., "89/100"). Nonetheless, it is obvious that a great deal of professional judgment and subjective evaluation necessarily goes into such an analysis. For example, the "environmental concerns" do not differentiate between a risk arising from toxicity, from volume, or from improper management, each of which has very different implications but all of which must be integrated by the analyst into a single number for that cell. Moreover, in a more detailed LCA, we would want to be able to weight different cells differently; a particular process might, for example, have significant implications for ecological systems, but very little for gaseous residues, which would incline us to weigh the ecological factor more heavily.

This technique, and indeed streamlined LCA methodologies generally, reinforce the principle that the purpose for which an LCA study is done should be considered in the choice of method. This matrix technology, for example, can provide useful information for a team seeking to build an environmentally preferable factory; it would not be rigorous enough to support legal or marketing claims to external communities or customers. In other words, it is the sort of study a firm would undertake for internal improvement, rather than external claims, especially if those claims are likely to be challenged by critical stakeholders. These limitations, however, are also strengths: Because there are obviously a limitless number of matrices that we could develop should we chose to do so, we can create custom analyses reflective of the particular circumstances, resources, and options that we face. The one element that should be included unless there are specific and compelling reasons not to do so is, of course, the lifecycle state approach: not only is this the heart of LCA, but most common failures to identify issues arise because the analysis has missed a critical lifecycle state. For example, early environmentalism tended to focus on manufacturing activities and related emissions, a partial and inadequate view that only later was superseded by the understanding that materials, products, and services have impacts across their lifecycles, and that such impacts could be far more important over time than those associated with manufacturing activities.

Finally, it is important to remember that these techniques tend to focus on environmental considerations. If they are to be used by the sustainable engineer, therefore, they either need to be modified to reflect social, health, safety, and other issues, or they need to be augmented by similar studies and information from those domains. Environmental issues are important, but in a sustainability analysis they are only a part of the suite of considerations, a point that must be reflected in tool selection, data generation, and decision processes.

7.5 SYSTEMS ENGINEERING

There are many methods, techniques, and fields of study in virtually all engineering disciplines that fall under the rubric of "systems engineering." In fact, virtually all engineering is systems engineering in some form or another, because we almost never run into an engineering problem completely bounded within itself. Solving a problem successfully almost always requires some understanding or explicit integration of systems state. Thus, for example, design tools for products, such as the House of Quality, are ways of bringing customer preferences, regulations, other technologies as appropriate, and so forth into the design process. A bridge is designed the way it is because it must handle a certain type and amount of traffic, and that traffic is also a technology system. (Building a bridge that was a foot too narrow for cars on a main road would be regarded as fairly stupid.)

Most systems engineering methods are highly quantitative, and will be introduced to the student in the appropriate disciplinary classes. Accordingly, this section will simply outline a typical systems engineering process, with the dual goals of familiarizing the student of sustainable engineering with the principle, and suggesting that, like industrial ecology, it can be a useful source of insights and procedures for sustainable engineering.

Although there is wide variation depending on application and situation, a useful schematic of systems engineering as a process might include the following steps:

1. ***Determine the goals of the system.*** This sounds easy, but in practice it can be quite complex. It requires that one first identify the communities and institutions that should be consulted regarding the particular system. The client will certainly be one important contributor, but especially where social or environmental systems are involved, other stakeholders will also be relevant (e.g., environmental NGOs, community activists, customers of the product or infrastructure). A second substep is to establish a dialog with each contributor so that agendas and goals can be defined. In some cases, it may well be the case that the explicit statements of particular stakeholders differ from their implicit agenda, and this needs to be part of the initial assessment. For example, an environmental group may say it doesn't believe that appropriate research has been done on a particular GMO technology, which establishes its explicit goal, but the implicit goal may be to stop all GMO technology development. In such a case, it may well be that there are no possible solutions that meet all goals, but there may be solutions that meet enough of them to be viable in the real world.
2. ***Establish criteria for ranking alternative solutions.*** This step involves the difficult task of converting the goals you have identified

into measurable quantities, so that you can both implement more formal design procedures, and know whether or not you're actually meeting the identified goals, and if so to what extent. All goals may not lend themselves to becoming metrics; during the process of determining metrics, "goals" may end up being reinterpreted as bounds, limits, or constraints.

3. ***Develop alternative solutions.*** This step is at least conceptually simple; it is where you use your professional training and experience, and a large dose of imagination and creativity, to develop solutions that may include technological, functional, and long-term structural alternatives. For example, an engineering team charged by the State of Arizona with building a replacement bridge in northern Arizona across the Colorado River onto Navajo Nation land was at one point in difficult negotiations, as the Navajo Nation was somewhat distrustful of the proposal. The solution developed included building a new bridge, which had a well-designed and aesthetically appealing, visitors' center with exhibits on Navajo culture on the Navajo side of the river as part of the installation. (In fact, the bridge is called the "Navajo Bridge.") It was an imaginative solution that met the concerns of the Navajo Nation while at the same time allowing the enhanced functionality of the new bridge that the state desired. Note also that a traditional engineering approach would not have regarded this project as one involving systems engineering: Building a bridge is, well, building a bridge. However, the cultural and political complications of this project were substantial, and a sustainable engineering approach—one that understood the challenge as involving not just engineering, but economics (Arizona's need for a new bridge) and social and cultural systems (not just the Navajo, but environmental organizations, were seriously concerned about the project)—was in fact necessary to get the bridge actually built. In many ways this sort of complexity is the rule, not the exception, in modern engineering (and a major reason for the rise of sustainable engineering).
4. ***Rank alternative solutions.*** Here you must include not just the usual technical and economic performance metrics, but a number of other considerations such as:
   - Effects on nonusers or minor stakeholders
   - Effects of piecemeal or incremental implementation
   - Impact on/coupling to existing systems
   - Sensitivity to parameter variations (look for robust solutions)
   - Ratification procedures (systems solutions almost always require political acceptability in real world)

5. ***Initiate action.*** This is generally self-explanatory. Any action should, of course, be measurable and connected back to the goals initially established. Moreover, in many cases, whatever action is taken may well be a continuing process rather than the end of a project. For example, the Navajo Bridge was a project with a beginning and an end, but if one is, for example, engineering the Everglades, there is no real end to the responsibility. Individual actions—e.g., taking down a particular levee on a drainage canal—may be bounded, but the overall project, to create an Everglades that meets the goals established with various stakeholders (protect avian biodiversity in the Everglades, support Florida agriculture, maintain water supplies for growing Florida communities), is an ongoing process requiring continuing dialog.
6. ***Iterate on both implementation and system response.*** It is not enough to begin well. With a complex system, it is important to dialog with the relevant elements, including how the system itself is changing over time, to know whether particular actions are having the desired effects. In this aspect, systems engineering is, in fact, an on-going learning process.
7. ***Design and manage governance as part of the ongoing iteration between system and design.*** This is a comprehensive activity that should be an explicit part of the system engineer's responsibility. Although many projects may have a relatively fixed and well-established governance system—say, project management guidelines for civil engineering of urban infrastructure projects—others, especially in the earth systems engineering and management arena, will require a much more fluid and evolving governance structure. For example, if you are one of the engineers charged with designing the water flow systems of the Everglades, you will be operating in a political and governance environment that is not only undefined, but constantly changing over time and with shifting priorities. In many cases, of course, the engineering team will only be a part of the governance regime, but especially given the increasing importance of technology in many of these systems, that does not mean the sustainable engineer can overlook this element of responsibility.

Systems engineering is thus an appropriate segue from the fairly applied and bounded world of specific materials, processes, and products with which much of industrial ecology, and tools such as DFE and LCA, deal, to the broader question of managing large and complex earth systems. Sustainable engineering draws on industrial ecology tools, systems engineering experience, and management innovations such as adaptive management and earth systems engineering and management, discussed in Chapters 13 and 14, to create new capabilities to address complex adaptive systems.

Exercises

7.1 You are a new design engineer for Widgets, Inc., having been hired for your training in sustainable engineering. You have been asked to select two materials for use in a new product line. One will be used in minor amounts, both in absolute terms and as an incremental increase in global demand for the material, as an additive to a stable glass matrix, while the second will be required in large volumes, both absolute and in terms of global demand, as a coating that you anticipate will shed over time with use. Both are somewhat toxic. From an industrial ecology perspective, what sorts of questions might you ask about these materials, and how would they be different for the two materials given their profile?

7.2 You are doing a sustainable engineering evaluation of the second material discussed in the first exercise above. How would you use an LCA as part of a sustainable engineering evaluation?

7.3 You are told that, on a life-cycle basis, battery-operated vehicles pose slightly higher risks to human health than vehicles powered by low-emission internal combustion engines, because of power plant emissions and the dangers associated with the disposal of lead-acid batteries. How would you go about verifying this claim?

Annotated Bibliography

1. There are a number of good resources available in the area of industrial ecology. Probably the best step up from this text is the T. E. Graedel and B. R. Allenby's 2010 text, *Industrial Ecology and Sustainable Engineering,* published by Prentice Hall. A more policy-oriented discussion that is somewhat dated but still relevant is B. R. Allenby, 1999, *Industrial Ecology: Policy Framework and Implementation* (Upper Saddle River, NJ: Prentice Hall). A good general handbook is R. U. Ayres and L. W. Ayres, eds., 2002, *A Handbook of Industrial Ecology* (Cheltenham, UK: Edward Elgar). And, of course, there is the original September 1989 article that started "modern" industrial ecology by R. A. Frosch and N. E. Gallopoulos, "Strategies for Manufacturing," *Scientific American* 261(3): 144–153.
2. One of the original works on DFE, useful because it shows how simple some of the early improvements really were (at least conceptually), is the 1992 compendium from the American Electronics Association (AEA), *The Hows and Whys of Design for the Environment: A Primer for Members of the American Electronics Association* (Washington, DC: AEA). In addition, every year since 1993 the IEEE has held the Annual International Symposium on Electronics and the Environment (currently renamed the International Symposium on Sustainable Systems and Technology),

and the *Proceedings* from those conferences is probably the best overall technical resource on LCAs of products and processes, DFE, and to some extent industrial ecology. An additional resource on DFE and LCA is J. Fiksel's 1996 *Design for Environment* (New York: McGraw Hill). For up-to-date information on LCA, the best thing to do is search the SETAC Web site (www.setac.org), or look at EPA's LCA Web sites (start with www.epa.gov/nrmrl/lcaccess/). The best systemic LCA program available on the web right now is the Carnegie Mellon Green Design EIO-LCA, available at www.eiolca.net. More broadly, there are a number of professional organizations, especially in Europe, that focus on LCA and similar issues; their Web sites and publications are good sources of state-of-the-art information.

3. Although the early plans are somewhat dated at this point, the integrated policy plans developed by The Netherlands over the past two decades are useful both for demonstrating how policy can integrate across social and environmental issues, and for the perspective they offer as a timeline of similar documents on the emergence of industrial ecology/sustainability practices and thinking. Begin with the 1989 document released by The Netherlands Department of Housing, Physical Planning and Environment, *National Environmental Policy Plan: To Choose or to Lose,* the subsequent 1990 *National Environmental Policy Plan Plus,* and the 1999 *Third National Environmental Policy Plan*, which can be found at www.un.org/esa/agenda21/natlinfo/action/netherla.htm (accessed January 2010). The 4th National Environmental Policy Plan was recently released, and is expected to provide guidance out to 2030 (the Plan is not available in English, but for the Dutch version see www.vrom.nl/Docs/internationaal/NMP4wwwengels.pdf, accessed January 2010).

CHAPTER

8 The Five Horsemen: Emerging Technologies

By the twentieth century, our numbers, our high-energy technologies, and our refined division of labor with its exchange economy made us capable of total transformation of any and all ecosystems. . . . In the twentieth century we became what most cultures long imagined us to be: lords of the biosphere.

—J. R. McNeill[1]

8.1 INTRODUCTION

Chapter 5 provided a glimpse of the power of technology systems in its discussion of the railroad, but we now face a wave consisting of not just one core technology, but five, the **Five Horsemen**: nanotechnology, biotechnology, robotics, information and communication technology (ICT), and applied cognitive science. (Cognitive science may sound like a science, but the development of technologies and fundamental scientific advances are so intertwined that it is a technology domain as well.) Moreover, these technologies are not just potent in themselves; they integrate in new and unpredictable ways across the technological frontier. Nanotechnology, for example, enables new kinds of ICT devices that, in

[1]J. R. McNeill, *Something New Under the Sun* (New York: W. W. Norton & Company, 2000), pp. 193–194.

turn, create the ability to build huge data systems in the biotechnology realm, as genomic and proteomic research explodes. On the one hand, these technologies are changing so rapidly, and the implications are so difficult to foresee at this time, that any discussion is necessarily incomplete and somewhat speculative. On the other hand, to speak of sustainable engineering without at least some idea of what's already coming down the pike is untenable: no engineer should be considered adequately educated, much less competent, in sustainable engineering unless that engineer has an idea of what emerging technologies are and what they mean.

Accordingly, we will cover several topics in this chapter. First, we'll define the relevant technologies and provide a little bit of background on their evolution, which, notably, is going exponential by many measurements for all of them. Second, we'll discuss some of the implications of emerging technologies in terms of a favorite subject for most people—which is course people. Although there are a lot of terms for this sort of thing, perhaps the best known is *transhumanism.* However, as this is a highly contested term culturally (see Box 8.1), we'll discuss the importance of the human becoming a design space and how that single development alone makes the future potentially very different from anything humans have known in the past. Heretofore, for example, engineering in general has been able to assume that "the human" is stable physically and mentally, and therefore concentrate on designs for systems, even very long-lived systems, that reflect this assumption. But today, some systems are being designed without that assumption. Some fighter plane designs, for instance, are now being designed from the beginning to operate in three potential modes: manned (the traditional mode); unmanned flown from a distance (the existing unmanned aerial vehicle [UAV] model, such as the Predator in Afghanistan); and directly coupled to the human brain, either as a single vehicle or as a grid of fighters.

An important theme of the discussion will be the importance of technological evolution to cultural power, especially now, when many nations—including the United States, China, Russia, Brazil, India, and perhaps others such as Indonesia and Mexico—realize that technological competence is not only a critical entry card to global status, but an important playing field on which national and cultural destinies will be determined. This is not just true in terms of achieving and maintaining global economic status, but in a purely military sense as well. One way, in fact, that the sustainable engineer can keep a close eye on emerging technologies as a whole is to track the technologies, investments, and research agendas of military establishments around the globe, an area of sufficient importance to deserve its own discussion in Chapter 11.

In this light, if the history of technological evolution is a warning, it is an entirely inadequate one for the wave bearing down on us. Technological change, as suggested by the example of the railroads, is always potent, but now we have the Five Horsemen. And because we are somewhat used to them, it is easy to ignore their profound importance.

BOX 8.1 Transhumanism

From one perspective, the integration of humans and their technology is old news—really old. A Neolithic human with a stone ax, and the knowledge to use it, is a different creature than one with neither. Today, none of us think twice about relatives with replacement joints, or showing up in glasses or, for that matter, drinking coffee for cognitive enhancement and using Google as a form of memory enhancement. In fact, the philosopher Andy Clark argues that humans have always been cyborgs, and that our major competitive advantage as a species lies in our brain's unique and innate ability to couple to external systems—including not just technologies, but social, economic, and information systems—so we can evolve integrated human/technological capabilities. This argument would lead to a conclusion that it is not a question of "trans" or "post" human, because being human has always meant being integrated with ambient technology. Rather, it is simply the much more rapid pace of technological evolution, and its increasingly profound effect on traditional ideas regarding what "the human" actually is, that have led people to create such concepts.

That said, one should recognize certain terms because they are in some sense already eyes of cultural hurricanes—*sustainability* is to some extent such a term, as is *transhumanism*, which, although it literally means extending the human, is also part of a cultural conflict. Thus, the World Transhumanist Association (the quoted wording below was subsequently modified and redefined, and the name of the organization changed to Humanity+ in 2008, in part perhaps to reduce the cultural conflict increasingly associated with the term—see http://humanityplus.org/) defined *transhumanism* on their Web site, www.transhumanism.org, as:

1. The intellectual and cultural movement that affirms the possibility and desirability of fundamentally improving the human condition through applied reason, especially by developing and making widely available technologies to eliminate aging and to greatly enhance human intellectual, physical, and psychological capacities.
2. The study of the ramifications, promises, and potential dangers of technologies that will enable us to overcome fundamental human limitations, and the related study of the ethical matters involved in developing and using such technologies.

From an analytical perspective, this definition—which as noted has been significantly changed on the Humanity+ Web site, identified a couple of important themes. Most important, by using terms such as *movement* and *study*, rather than referring to particular physical states or integration of biological humanity and technology, the definition signaled that *transhumanism* is a cultural and intellectual concept, rather than just an engineering framework. In addition, the reference to "applied reason" should also be a flag, because if the human as design space means anything, it means one is dealing with a complex system, where traditional methods of analysis, while important, are unlikely to be completely adequate to design or

manage the system as a whole. This becomes important when considering the language about "fundamentally improving" the human condition; it is, after all, also possible that unpredictable technological evolution, integrated into the human, will transcend or even render obsolete important aspects of what we might value as human, at least currently. Many of us may agree, for example, that all else equal, enhancing cognitive abilities or reducing pain and suffering is desirable, but the technologies that can do so are potent enough that they will likely have other systemic effects—might they make us less prone to take risks, for instance, or reduce the barriers to war in some ways? For example, if radical life extension is achieved, an early death is not just the loss of a few decades of life, but of perhaps a century of life. Does one's risk calculus change?

Not surprisingly, then, "transhumanism" has in its turn created a strong opposition, especially among conservative and fundamentalist religious thinkers. Some of these individuals argue that even medical progress has reached a point where it is increasingly inequitable, futile, or misguided, and that pain and suffering are an important part of the human experience. (Those of less religious bent may remember in the movie *The Matrix* that the first virtual reality created by the machines was made so that humans could always be happy, but that reality had to be redesigned because it turned out that humans couldn't adapt to happiness on a permanent basis.) Some even argue that human enhancement constitutes blasphemy because it involves deliberate redesign of humans who are themselves created in the image of God, and is a primordial sin against the order that God has established, the Great Chain of Being that gives everyone, and everything, its proper place.

The "**Great Chain of Being**" is an important concept for anyone working on technologies that challenge current ideas about what it means to be human, because it is often the unrecognized and unconscious source of the feeling that any change in humans is "playing God." The Great Chain is a conceptual framework of the Universe perfected in the Christian medieval period in Europe with roots in classic Greek thought, especially Plato and his allegory of the Cave, for example, with its elevation of idea over observation, and Aristotle's *De Anima*, which differentiates plants, beasts, and humans based on the functionality exhibited by their souls—plants having less than beasts, which in turn have less than humans. The Great Chain defines a structured hierarchy, with pure spirit and perfection—God—at the top, and pure matter and imperfection—rocks and other materials—at the bottom. In between, in order, come all things; angels are next to God, while plants, then beasts, are above matter. Humans are at the fulcrum point, for they are both spirit and matter. This framework is important for several reasons. Politically, it served to justify the strong hierarchical structures that characterized society and the Church at that time. But because the Great Chain of Being is given and fixed, not a matter of choice, changing it as transhumanism implies is a violation of the basic order of the Universe and God's will, and is thus blasphemous. Many current Christians, although they may not be fundamentalists in the accepted sense, still implicitly accept the validity and reality of the Great Chain.

From the perspective of the sustainable engineer, this little digression into transhumanism and theology has some important implications. First, be careful of language. To those who are working in areas of emerging technologies and human systems, *transhumanism* may sound like simply another way of talking about human enhancement, but it is a culturally and politically loaded term. Use it only if you understand it and how it may affect your target audience.

Second, the integration of humans and their technological systems may, upon reflection, strike most engineers as a statement of the obvious. But it is important to remember that engineers, and trained technocrats generally, are a minority in any society, and that there are other groups in society for whom such integration, especially as it becomes more obvious and disruptive of existing cultural patterns, will be perceived as threatening, if not downright evil. Particularly as technology becomes more complex and difficult to understand, those who are increasingly powerless, whether because of lack of education or cultural weaknesses in such a competition, are liable to strike back through fundamentalism. It is not improbable that, as such groups begin to understand that redesign of the human is already here, and that such redesign is strongly correlated with cultural and economic power, the reaction that has started will intensify, perhaps dramatically. It bears remembering that the invention of printing enabled the spread of literacy and democratization of the Bible—thence was the technological foundation required for the Reformation. Printing, then, may be thought of in some way as redesigning each individual to be able to talk directly to God, without the Church as necessary intermediary—the result was 150 years of religious conflict across all of Europe. You may think that your work, be it on pharmaceuticals that reduce fear or ease the psychological effects of trauma, or on information systems that feed ever more directly into brains, are an obvious benefit to society, but, remember, there will be others that disagree, perhaps strongly.

Moreover, your job as an engineer is to solve problems. Some technologies have stronger social and cultural implications than others, however, and it may fall to you, as it did to Monsanto (see Box 2.1), to address cultural, social, or environmental issues far beyond those you expect. It may also fall to you to work with, and integrate into your solutions, worldviews that you find different, perhaps even mutually exclusive. That, also, is part of sustainable engineering and of the sustainable engineer as leader. Certainly the shift of the human from something we assume as stable, to a design space, will represent such a challenge.

8.2 THE FIVE HORSEMEN

In considering the most potent of the emerging technologies, let us begin with nanotechnology. This technology system extends human will and design to the atomic level; it is in some ways the culmination of the human project to gain ascendancy over physicality. Perhaps because of this long history, however, the term itself is somewhat ambiguous; after all, chemistry has long dealt with

phenomenon at the nanoscale (a nanometer is one thousand millionth of a meter), and the scale of many electronics devices, such as circuits on a chip, has long been small enough to fit within some definitions. But most definitions emphasize three important elements of nanotechnology:

1. Deliberate manipulation of materials, machines, or constructed devices at the atomic, molecular, or nanoparticle scale
2. A spatial scale reference, which is usually considered as under 100 nanometers
3. Design, purpose and function that take advantage of, and focus on, new and novel properties arising from the small scale of the artifact or material particle

These characterizations are useful starting points, but are still not definitive—after all, bioengineering occurs at the nanoscale in many cases. Is biotechnology, then, a subset of nanotechnology? To further complicate analysis, nanotechnology is both an area of technological research and design on its own, and an enabling technology across much of the existing scientific and technological frontier. Moreover, many people conflate "nanotechnology" with "nanomaterials"; while the latter is obviously a part of the former, the two are not the same. Accordingly, some have projected potential pathways for nanotechnology development; a typical one might include four sequential stages:

First Stage: nanomaterials in various configurations

Second Stage: simple but active multitasking nanostructures, such as actuators, drug delivery devices, sensors and the like

Third Stage: nanostructure systems containing thousands of interacting components

Fourth Stage: integrated nanosystems, with complexity at the scale of the human cell

Looking at existing data, it is fairly clear that we are solidly in the First Stage: nanomaterials are in the takeoff stage of their exponential growth phase, as Figure 8.1 illustrates. Although bulk applications make up the preponderance of production by weight, it is likely that, at least initially, health care and electronics will be major drivers of economic value. (Many low-value, high-volume nanomaterials, such as titanium dioxide, flow into structural applications, whereas the high-value nanomaterials will flow towards health care and electronics.) By contrast, in 2009 at least, there are few operational active multitasking nanostructures, although the research literature, especially regarding drug delivery devices, is rapidly growing richer and more sophisticated.

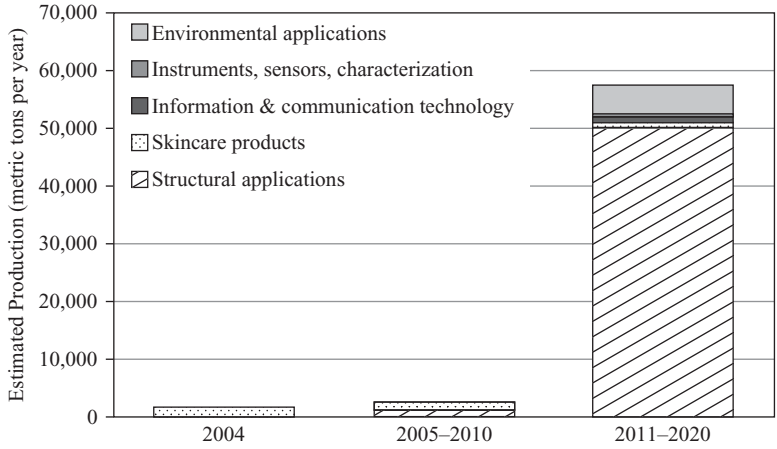


FIGURE 8.1 Estimated Nanomaterial Production. This figure shows a projection of anticipated growth in nanomaterial production, and in particular suggests that the exponential takeoff characteristic of technology systems in the early stages of their growth may be near if not already in progress. Like all projections, this one necessarily relies on a number of assumptions, and is thus best regarded as a likely scenario rather than something that will necessarily come to pass. It also means that one should be aware of the institution behind the projection, and whether it has an interest in particular outcomes. In this case, the Royal Society is a well-respected scientific and technological professional organization with no commercial interest in nanotechnology.

(Data from: Royal Society and Royal Academy of Engineering. "Nanoscience and Nanotechnologies: Opportunities and Uncertainties." http://www.nanotec.org.uk/finalReport.htm: accessed October 30, 2009. Printed with permission.)

Biotechnology extends human domination across the biosphere (hence the quote from J. R. McNeill, an environmental historian, that begins this chapter). As with the other emerging technologies, technological evolution in biotech is accelerating dramatically. This is illustrated in Figure 8.2, which shows that technological capability as measured in synthesis productivity and synthetic DNA length is advancing exponentially (note the vertical axis in the figure is logarithmic), while the costs of gene synthesis and cost per base sequenced are falling equally rapidly. These trends are not only powerful in themselves, but they underpin equally dramatic acceleration of the human ability to design biological systems, including human systems, at the level of metabolic pathway, cell, organ, and organism.

Thus, the human genome has been mapped, and extensive research efforts to extend that mapping to include proteins and understand gene expression continue. Similar information is being developed on a rapidly increasing number of bacteria, yeast, plants, and other mammals. At the organism level, genetic engineering of species is proceeding apace: to increase yield; reduce pesticide consumption; reduce demand for land for agriculture

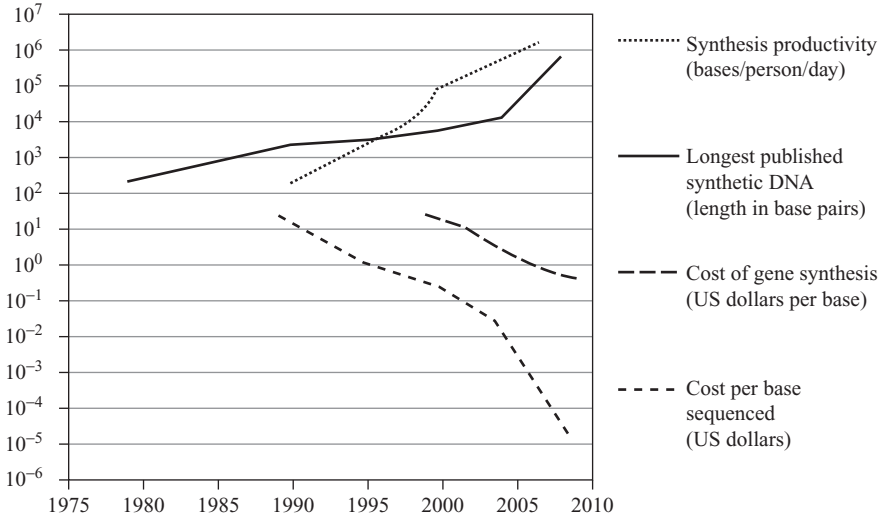


FIGURE 8.2 Carlson Curves. This graph illustrates the acceleration in biotech capability, an exponential growth process also evident in the other Five Horsemen. This means that growth is both rapid and that the rate of growth is accelerating. Note that the vertical axis is marked off by powers of ten, not arithmetically. This has the effect of turning exponential growth into a straight line and allowing data that covers many orders of magnitude to be displayed on a single graph. Compare this figure to Figure 8.5, which is similarly constructed.

(Data from: Carlson, R. "Synthesis." http://www.synthesis.cc/: accessed September 18, 2009. Reprinted with permission.)

and energy production; enable plant growth under saline conditions, thereby conserving fresh water resources: produce new drugs; reduce disease; increase hardiness; and support a healthier human diet. Indeed, the development of designer foods, engineered to produce either nutrition or health benefits, is itself growing exponentially, a trend that will accelerate as the technology to produce meat in factories, rather than raise animals for meat, is developed and deployed. (Nutritional benefits arise when food is designed to contain more minerals, vitamins, helpful bacteria, or other additives than it otherwise would; health benefits, a more controversial area, arise when such additives are connected by the manufacturer to specific health benefits—for example, when a product that is enhanced by adding calcium is claimed to "promote strong bones.")

As a result of such work, a new field called "**synthetic biology**,"—or less formally, "plug and play biology"—is developing. Synthetic biology merges engineering with biology by creating standard biological components that can be mixed and matched in target organisms to provide desired functions. This allows researchers to create biological components, circuits, and potentially replicating organisms from scratch, and design extensions beyond existing biological systems in new and completely anthropogenic ways, such as creating life based on different genetic codes than those found in the wild. Massachusetts Institute of Technology, for example, has established a Registry

of Standard Biological Parts ("BioBricks") that can be ordered and plugged into cells, just like electronic components. Less playful examples of synthetic biology include controversial breakthroughs such as rebuilding the 1918 flu virus, and the polio virus, from scratch. Dr. Venter announced in 2010 that he had built a bacterial genome from scratch and implanted it in a recipient cell from which the indigenous DNA had been removed, where the constructed genome enabled the cell to successfully replicate. Other researchers have engineered the genes of *Escherichia coli* to incorporate a twenty-first amino acid, opening up an option space for design of biological organisms that has been unavailable for billions of years and leading to the design of life forms that are not merely extensions of evolved biodiversity, but never would have existed at all but for human design choices.

The sustainable engineer should note well the enormous implications of the kind of technological evolution represented by these developments in biotechnology. First, synthetic biology dramatically changes what biology is. For example, a standard and very important concept in biology is the species, but once humans begin to shift genes between organisms without regard to their taxonomic status, the idea of species becomes increasingly obsolete: Biology becomes a series of design choices, not taxonomic relationships. Second, synthetic biology dramatically increases the option spaces for design of life, not just by offering opportunities to put jellyfish genes in mammals (for instance), but by creating entirely new opportunities to design entities from the genomic material up. In short, life forms can be designed that could never evolve "naturally" on our planet. Because synthetic biology involves design choices towards ends such as efficiency or enhanced aesthetic appeal, it creates an entirely different world than that to which the nondirected process of evolution would lead: Synthetic biology is designed to do something, or to be something, it doesn't just evolve. Third, it changes what biology is. Biology is no longer a "natural science," rather, it increasingly becomes an engineering discipline, driven by economics, culture, and social needs. This is a shift that is not only invisible to those working in much of biology today, but that raises difficult questions regarding whose worldview should be privileged above others by policy makers. The "biodiversity crisis" is seen very differently, for example, by those in conservation biology than it is by those who are creating new life forms in their laboratory (see Box 8.2).

Robotics is a more applied technology system than nanotechnology or biotechnology; indeed, it draws from both of them (in the case of experimental robots that use a neuronal mass from rats as part of their control and behavioral system, for example). Because of their economic role, use of robots has expanded, but at a slower pace than other technologies, in part because robots are generally used in manufacturing and other applications where economic, social, and institutional factors dampen dramatic shifts in short time

BOX 8.2 Who Defines the Playing Field? Conservation Biology and the Biodiversity Crisis

For many people, science is seen as a purely factual and objective activity, with results founded on experiments that can be duplicated if necessary to confirm assertions. But this reliance on the scientific method, which is indeed powerful, should not lead the sustainable engineer to overlook the serious difficulties that arise when periods of rapid change pit different worldviews against each other, each supported by scientific argumentation and disciplinary structures. Consider, for example, an issue of deep concern to many environmentalists, sustainability advocates, and conservation biologists: the "biodiversity crisis." In a nutshell, the argument, supported by data developed by conservation biologists, is that human activity, especially agriculture and the widespread changes in water management and land use that accompany it, has precipitated a dramatic and irreversible increase in species extinction. Others have begun to argue that what appears to be an extinction crisis if only evolutionary biodiversity is considered, may in fact simply be a cusp as the source of biodiversity shifts from natural evolution to human design, and that reductions in information content in biological systems resulting from extinctions of evolved life forms may be balanced by creation of new information content as new bioengineered life forms are designed and commercialized. In short, evolved biodiversity may be shrinking, but the total information content of the biosphere is perhaps not shrinking so much, and may even be growing, as synthetic biology continues to rapidly develop.

Whether this dynamic is indeed at work is obviously in part a very complicated factual question (how does one quantify the information in biological systems, especially given scale differences as one moves from the genome to the ecosystem?). But this dynamic is more than that: It is a cultural and a political power question as well. In this regard, perhaps the most interesting question is: What group gets to define the cultural construct of biodiversity, and why does it matter? As it turns out, it matters a lot. The community that controls the definition controls social perception of the phenomenon, with significant policy implications.

To begin with, it is apparent that "biodiversity" is not a factual observation, but a cultural construction. One way to construct it—the way that we generally have for reasons of history and cultural belief—is by considering only biodiversity that arises from evolutionary processes, in which case loss of "traditional" species equates to loss of biodiversity. Alternatively, we can consider only designed biodiversity, in which case gains in constructed forms of life, such as GMOs or engineered bacteria, constitute gains in biodiversity. Or we could consider overall information content of biological systems of all sorts as biodiversity, in which case no one knows whether it's increasing or decreasing. How to measure the overall information content of the biosphere is quite an interesting question in itself, especially in a period when the biosphere is shifting from an external system to be observed, to a human system to be designed. More fundamentally, we really don't have a very good idea as to how many species there are to begin with.

What community gets rights to define the cultural construct of biodiversity matters because the definition of the term in large part bounds public perceptions, and thus potential policy responses. *Biodiversity* traditionally has been defined by the conservation biology community, which by self-selection, focus, and training has a strong predisposition to perceive losses in evolutionary biodiversity, and little experience or interest in understanding designed biodiversity at all, except perhaps as it overlaps "natural" systems. (A major reason is that designed biodiversity tends to be found in industrial and agricultural systems, as opposed to "natural" environments.) This definition accordingly is based on the construct of "species" generated by evolved biology, and given the continued development of the anthropogenic Earth, necessarily implies a "crisis" in biodiversity, requiring in turn strong policy measures to "preserve biodiversity."

For their part, synthetic biologists have had little incentive, or disciplinary pressure, to consider issues of biodiversity except defensively (that is, showing that their products will not adversely affect "natural" biological systems). Moreover, no scientific community has an interest at present in understanding biodiversity in the fully integrated sense, including both evolved and designed biological systems. This is primarily because the conservation biology concept of biodiversity is so dominant, but even if there were anyone interested in such a reconstruction of the concept, definitional and measurement issues are daunting. After all, "species" is a convenient measure of biological information content even if it is partial and increasingly misleading, whereas trying to understand the relevant measures of information of designed biodiversity is quite difficult. The end result is that it is the community that has the authority to construct the concept "biodiversity" simultaneously constructs a "crisis" and, in turn, policy responses to the perceived crisis. In other words, the state of crisis is inherent not (necessarily) in the data, but in the definition of the concept. None of this is to suggest that the changes in evolved biodiversity should not be of serious concern; it does, however, suggest that a sustainable engineering approach might want to frame the system more comprehensively than the current major disciplinary contributors to the dialog. Doing so might lead to important questions, such as what the implications are for the stability and resiliency of Earth systems if evolved biodiversity is replaced by designed biodiversity at more than minimal scale.

More generally, for the sustainable engineer, this discussion emphasizes the importance of understanding what community lies behind, and thus defines, concepts such as "sustainability" and "biodiversity." Such communities not only explore facts, but create cultural constructs that reflect far more than facts, because they are the structures that determine which facts "matter"—that is, become part of public dialogs and therefore become part of public policy responses. This in turn is important because most cultural constructs change existing distributions of costs and benefits (otherwise, of course, they would not be needed), and become embedded in designs and technologies. Thus, the community that creates such a construct that, in short, is able to be the "voice" on a particular issue, is able to redistribute costs and benefits across society in line with their belief systems.

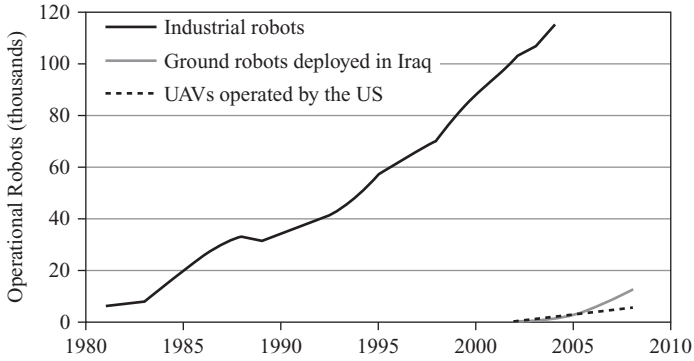


FIGURE 8.3 Robot Deployment by the United States. This graph illustrates the rapid uptake of robots in the United States. Industrial robots dramatically outnumber military robots, but the rate of growth in military robots, in terms of both numbers and, more subtly, in capability, is noteworthy.

(Data from: Singer, P. W. "Robots at War: The New Battlefield." http://www.wilsoncenter.org/index.cfm?fuseaction=wq.essay&essay_id=496613: accessed September 20, 2009; United Nations & International Federation of Robotics. *World Industrial Robots.* New York: United Nations, 1996; United Nations & International Federation of Robotics. *World Robotics*. New York: United Nations, 2005.)

periods. What is considered a "robot" also varies over time: older machine tools in some ways were robots, but they had nowhere near the versatility and capability of newer industrial robots. Figure 8.3 illustrates continued strong demand for industrial robots in the United States, but not the accelerating exponential growth seen in the cases of other technologies. Figure 8.4 illustrates

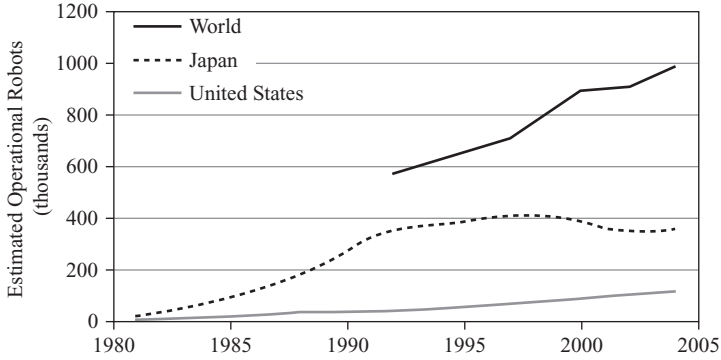


FIGURE 8.4 Industrial Robot Deployment. While use of industrial robots in the United States and Japan has flattened somewhat, growth continues strongly across the globe as a whole. Note that both Figures 8.3 and 8.4 display linear growth as opposed exponential or geometric.

(Data from: United Nations & International Federation of Robotics. *World Industrial Robots.* New York: United Nations, 1996; United Nations & International Federation of Robotics. *World Robotics*. New York: United Nations, 2005.)

a similar pattern for the world. These numbers must be taken with a degree of caution, however, in that robots used for domestic purposes by individual customers are not included; in Japan especially, much of the recent work in robotics has focused on developing robots that can substitute for humans in fields such as elder care, reflecting the coming demographic crisis as the ratio of working-age individuals to seniors in Japan continues to shift. In this, as in military affairs, robots represent state-of-the-art substitution of capital for labor in production processes—which means there is less need for expensive workers in factories, and less need for war fighters in harm's way in combat.

One of the keys to understanding the implications of robotics as an emerging technology is the continued integration of robotic capability and biological systems. This not only includes experimental attempts to build neurons into robotic control systems, but also experiments aimed at directly connecting human brain activity to robotic functions, from moving robotic limbs and enabling speaking systems for paralyzed individuals, to directly wiring human beings into major mechanical systems (see Box 8.3). Although such large-scale integrated human/mechanical systems are not yet feasible, smaller scale experiments aimed at simplifying the **computer–brain interface** (CBI) are proliferating. For example, helmets that enable gamers to interact mentally with their avatars in virtual realities and children to mentally control simple toys, such as balls suspended on air columns, have already been commercialized. Given experimental results where researchers are able to determine through analysis of electrical patterns of brain emissions what noun is being thought of, and even follow images to some degree, the currently available applications may be precursors of far more powerful and socially challenging technologies (for instance, the ability to scan people's cognitive activity without their being aware of being scanned). Hybrid systems such as the helmet combine different emerging technology systems—in this case, robotics, ICT, and cognitive science—and warn against becoming too firmly attached to any particular way of bounding "different" areas of technology, as the truly interesting and challenging technological developments usually represent such integrations.

Chapter 2 alluded to the explosion of information in recent years and its implications for increased social and cultural complexity, a point that affirms the importance of ICT as an emerging technology and as an increasingly critical infrastructure across the board. Indeed, the underlying technology continues to change at a pace that recently is outrunning the ability of firms and the economy to adapt. This is occurring at the level of individual technologies as well as in services such as social networking. In chip design, for example, Moore's Law, a relationship first stated by Gordon E. Moore, one of Intel's founders in 1965, suggested that the number of integrated circuits per chip

BOX 8.3 Integrating the Brain with the Machine

Humans in general and their brains in particular have always been somewhat integrated with the machines around them. After all, without a brain, a hammer or a saw has little obvious application, and writing is unintelligible scribbles; likewise, a society that sees millions of automobiles running around its roads causing relatively little damage to each other certainly has a brain machine interface of sorts working. What is of interest, therefore, is the immediacy and the bandwidth of the connectivity and the different performance characteristics of the integrated techno-human system.

We can begin by differentiating among three types of links. The mechanical and sensory links between a device and a person's body are the primary coupling today, including everything from the plumber and her tools to the gamer moving avatars around a game space with specialized controls. Automobiles, ships, airplanes—all involve various forms of visual and audio metering, tactile coupling to various control and performance mechanisms, and the like. Indeed, by this perspective a Predator pilot guiding his unmanned aerial vehicle (UAV) over the battlefield is simply a sophisticated version of the tool-wielding Neolithic hunter.

For purposes of understanding the modern challenges, however, it is the search for more direct connections between central nervous system (CNS) activity and the device itself that are of interest. These can be of two generic types. One involves devices placed on the skin or worn as a helmet, which detect and interpret electrical signals and make changes in the immediate environment as required. Wearable helmets that enable gamers to interact directly with their avatar is an example. This is already a well-explored category: The most obvious example is probably the cochlear implant, where an external microphone records an acoustic signal and sends it to a computer located outside the skull, which then turns it into an electrical impulse and sends it to an electrode implanted in the cochlea in the inner ear. As of 2009, there were at least 30,000 such devices already in use around the world.

The second type also detects electrical signals from the CNS and the brain, and causes appropriate changes in the environment, but it involves invasive placement of suitable electronics in the brain itself. Examples might include wearable prosthetics driven by motor signals recorded in the brain, then transmitted to a computing unit implanted subcutaneously; chips embedded in beetles or moth larvae, which then pupate into a moth that can be guided by signals sent through the chip; or even the famous experiments where a monkey with a chip in its brain moved a remote arm via a direct wireless connection and subsequently from its home base at Duke University in the United States kept a robot in Japan jogging remotely.

Each type raises its own issues. Because the linkage between the Predator in the theatre of battle and its controller in Utah requires a number of physical interfaces—the Predator pilot must work his joystick, and that mechanism's activities must be translated into operational signals for the Predator—it is more subject to upset

than a direct neural connection. Wearable devices must necessarily deal with signal degradation as they travel from inside the brain to outside the skull; they thus may have trouble deciphering complex signals in a very noisy environment. Any device directly implanted in the body or the brain raises the risk level of the system. While there is substantial research going on in invasive devices, there is some reticence involved unless the risks of an invasive device are justified by benefits for the individual involved. This is especially true where the device is to be inserted directly into the brain itself: the brain is protected by the so-called blood–brain barrier, and anything that penetrates that creates a risk of infection or damage that needs to be carefully considered.

Many of the issues suggested by machine–brain interfaces are not entirely new, but there are some that are potentially quite revolutionary. For example, if one can design a wearable helmet that can pull images, words, and letters from the brain (research intended to lead to such a device has already been funded) one is on the path to developing the capability to pull thoughts from the brain without permission and without an individual's knowledge. The implications of such a technology, coupled with widespread adoption of the "telepathic helmet," would challenge most existing social, cultural, and even religious norms. It would certainly radically affect any remaining idea of privacy: Even in totalitarian governments, thoughts have heretofore been unreachable directly. Or again, coupling major weapon systems directly into war fighters' brains could, by removing the individual from the theater, dramatically reduce casualty rates—thus, perhaps, making violent resolution of political conflicts more likely and reducing the barrier to a country's going to war. As individuals got used to having such large technology systems integrated into their body image, it would also radically affect their sense of self and perhaps their sense of community with other humans, who may themselves be enhanced in other ways.

would double approximately every two years. Figure 8.5 shows that this, in fact, has occurred.

The implications of such a consistent and rapid doubling of a foundational technological capacity are truly unprecedented—an inexpensive sedan today has more computing power than the *Apollo* spacecraft when it landed on the moon. But it is not just chips. Virtually every other index of information technology has shown this exponential growth rate: pixels per dollar, Internet information traffic, and information storage capacity per hard disk (illustrated in Figure 8.6). Perhaps most important from an application perspective, the computing performance per unit cost has been improving exponentially as well, also doubling approximately every two years.

Note that if this exponential rate of growth had occurred in only one technology, such as hard disk storage capacity, eventually the lagging

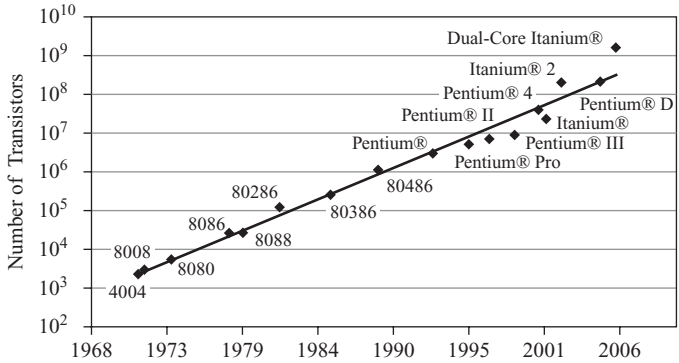


FIGURE 8.5 Transistors per Intel Microprocessor Chip. This graph shows that the number of transistors per chip has doubled roughly every two years, in keeping with Moore's Law (a prediction that has functioned primarily as a road map and goad to technological evolution).

(Data from: Intel Corporation. "Microprocessor Quick Reference Guide." http://www.intel.com/pressroom/kits/quickreffam.htm: accessed August 27, 2009.)

effects of coupled technologies like processors and bandwidth capacity would have limited the benefits, but because all the critical coupled technologies are evolving according to a similar exponential pattern, some systemic coherence in technological evolution can be achieved. Nor is this by accident. The electronics industry relies heavily on "road maps" and

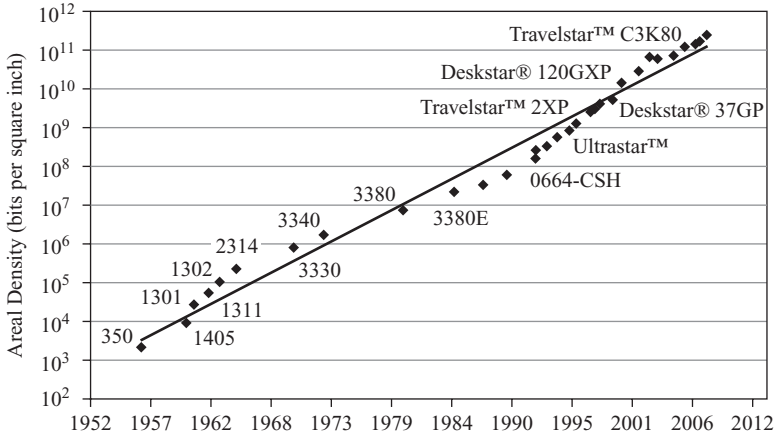


FIGURE 8.6 IBM and Hitachi Magnetic Disk Information Storage Density. This figure illustrates the Moore's Law growth of hard disk storage capacity along an exponential trend line. (Hitachi Global Storage Technologies emerged in 2003 from a combination of IBM and Hitachi's storage technology businesses.)

(Data from: Hitachi Global Storage Technologies and IBM.)

extensive road-mapping conferences and meetings, which work to assure that goals, technological developments, and interfaces across these complicated technologies are aligned. The sustainable engineer should be sensitive to cultural institutions, such as Moore's Law, and sectoral activities such as road mapping, that enable such coordination and guide technological evolution. The engineer should also be well aware of antitrust and other legal, regulatory, and commercial constraints and obligations that are integral to such processes.

From an ICT services perspective, technological change is also dramatic. One example is the demise of traditional newspapers, long expected to be under threat from Internet blogs and news sources, which are suddenly in serious and immediate jeopardy in many places. Although some may reinvent themselves in dual online/print format with more emphasis on local news coverage, and specialty niche media properties (e.g., the *Wall Street Journal* or *Financial Times*) should do well, newspapers as they now exist increasingly appear to be an obsolete communication technology. Wireline telephone companies in developed countries are facing similar challenges. In 2009, customers were abandoning landlines at a rate of 700,000 per month in the United States; in 2005, a little over 7% of all American households were cell phone only, but by the end of 2008, over 20% were. In 2008, the number of cell phone only families exceeded for the first time the number of landline-only families. Moreover, it is worth remembering that changes in these technologies have significant social and economic ramifications. For example, in the United States and elsewhere, it has for many years been government policy that landline telephone service subsidize rural telephone service. If that policy begins to fail because of falling landline revenues available for subsidy, then rural areas might become even more depopulated than they are now, because communication services are often necessary for any sort of employment in today's economy. That becomes another driver for accelerating urbanization, a trend we have already discussed.

These are massive changes in fundamental information technologies; if a student were curious about why such numbers might matter to him or her, just remember the number of telecommunications engineers who lose their jobs as these shifts occur. For example, despite robust growth in wireless, AT&T was shedding jobs at a rate of about 10,000 a quarter during 2010, primarily because its landline business was shrinking so rapidly. Other landline companies were behaving similarly. You lose track of these sorts of trends at your peril. Unlike jobs that fade because of recession only to come back again, the telephone engineers who lose jobs are unlikely to find anything like them in the future. The technology they know is dead. Indeed, one incentive for developing the broader perspective

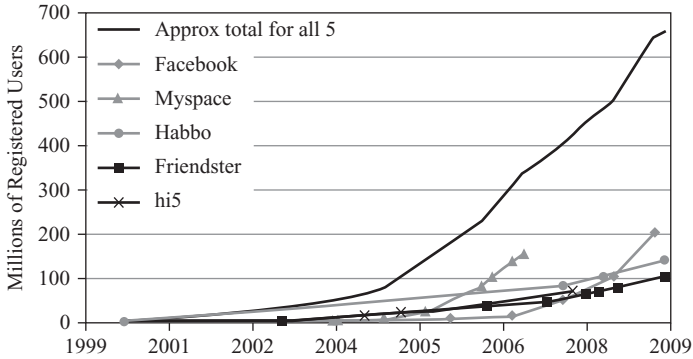


FIGURE 8.7 Participation in the Five Largest Online Social Networks. This graph illustrates the dramatic growth of five social networking sites: Facebook, which is a general social network with users around the world (ages 13 and up); Myspace, which is an entertainment-focused network (music, videos, etc., for ages 13 and up); Habbo, a network for teens with a global user base (ages 13 and up); Friendster, a network popular in Southeast Asia but not widely used in the West (and adults only, ages 18 and up); and hi5, another network popular outside the United States (ages 13 and up).

(Data from: Facebook, Friendster, hi5 Networks, and Sulake Corporation.)

of a sustainable engineer is to develop the flexibility and general skills to if possible avoid a similar fate in a future marked by even more rapid technological upheaval.

But it is not just the sheer volume of information, and virtually all of the underlying technology, that is changing. Both the social and cultural dimensions of information transfer are changing in rapid and unpredictable ways. One example of this is the dramatic and rapid spread of social networking technology, as illustrated in Figure 8.7. It is hard to believe, but the ubiquitous social networks that are already such a fixture for "digital natives"—people who have grown up with the Internet, virtual reality, and social networking technology—are post-2000 innovations. Facebook, one of the largest at this point, was only introduced in 2004 and passed its 500 millionth account in the first quarter of 2010; Myspace started in 2003. Virtual realities are similarly new. Second Life, perhaps the most popular virtual reality environment, was launched in 2003. (The other category of virtual reality is the massive multiplayer online game, or MMOG, featuring such familiar titles as Eve Online, introduced in May 2003, and World of Warcraft, currently the most popular, introduced in 2001 and by 2009 inhabited by over 11 million monthly subscribers.) The social and cultural implications of these technology systems are yet to be understood, and in all likelihood are still evolving themselves, but they will be profound.

More broadly, culture itself is an information origination, communication, and storage system, and it thus changes as ICT changes. At one level, that is reflected in the rise of the networked digital native and the gamer; at a different, perhaps more profound level, it is reflected in changes in the patterns that we use to think about our world. It is well established that people perceive their world in large part through cultural constructs, and that cultural constructs change over time. Thus, for example, Europeans settling in the New World because of their Christian culture tended to perceive "wilderness" as evil and saw their moral duty as "taming the wilderness." So John Quincy Adams was well within his culture when he called upon Americans "to make the wilderness blossom as the rose, to establish laws, to increase, multiply, and subdue the earth, which we are commanded to do by the first behest of the God Almighty." But now, of course, "wilderness" is conceived of as a great good—indeed, for many environmentalists it has assumed the role of the Sacred that in medieval times was ascribed to God in Europe. It is not that either is "right" or "wrong" so much as it indicates how powerful and how necessary such cultural constructs are. Indeed, as Box 8.4 points out, most ages have a set of basic interrelated metaphors that enable people to frame their complicated and changing world in ways that make it understandable. Beginning with Newton, when physics became the dominant science, humans tended to see their universe, and their world, as mechanism; now, in the age of environmentalism, we tend to see it in terms of ecosystems and sustainability.

The ever more rapid revolution in ICT also accelerates change in cultural constructs and underlying social patterns. For example, television and now the Internet are subtly changing the way that people understand and relate to cultural constructs such as "environment," "nature," and "wilderness." The surface effect, of course, involves substituting the passive consumption of information for actually experiencing the outdoors. But it goes deeper. People begin to understand "nature" as it appears on television: Animals are always active, scary, and challenging. This creates a growing gap between the concept of "nature" that people form from their technologies and their increasingly urban experience, and what's actually out there. More subtly, it gets people used to the idea that "nature" is something to be consumed: watched, bought, and sold. The end result is that more and more people are exposed to "nature" primarily on cable systems and shows competing for their attention, in upscale stores in fancy malls, or to "wild animals" in parks and "sea worlds," carefully designed by corporate sponsors to present a particular concept of wildness.

The underlying effect is both overlooked and profound. "Nature" shifts from being a system outside the human, something that we exist in,

BOX 8.4 Using Metaphors and Analogies

A sustainable engineer must be able to communicate as well as be technically competent. This requires practicing a different set of skills than just learning quantitative analysis, with the answers in the back of the book. It requires knowing a little bit about how people generally view their world, and how to express ideas and issues in such a way that they can understand it even if they lack technical training. In this regard, it is important to understand the significance of metaphors and analogies, and how to work with them.

As the text points out, humans throughout history have been prone to create archetypal metaphors for themselves. For the Confucian cultures, a well organized bureaucratic structure, with delineated authority and responsibility in well-defined hierarchies, is a powerful image. Indeed, some Chinese folk conceptualizations of heaven have a bureaucracy that is quite similar to the Confucian ideal. Similarly, the Greek philosopher Plato saw the three-part human soul that he postulated reflected in the structure of his ideal state. And with the Enlightenment and the dramatic successes of science in explaining phenomenon that had heretofore been explored through Biblical authority, the metaphor became the Universe (and things in it, such as human beings) as mechanism. Currently, reflecting the cultural power of environmentalism, the reigning metaphor is ecological systems: urban planners talk of ecocities and firms talk of their ecotechnologies. Indeed, entire fields may be extended metaphors; the new discipline of "industrial ecology" reflects in its name the idea that industrial systems can be understood and optimized by considering them as if they were ecosystems.

Remember that part of your job as a sustainable engineer is to provide leadership and advice on questions of technology, and help develop and support rational and ethical industrial practices and public policy. This requires good communications skills: in this regard, metaphors, analogies, and even stories, are useful because they help make complex ideas approachable. For example, using the template of an ecosystem to try to make sense out of highly complex industrial systems, and think about ways in which they could be made more environmentally efficient, is not a bad idea. But they can become destructive if they become so powerful that they blind one to differences as well. For instance, the idea of an ecocity, to the extent that it implies a city that uses resources more efficiently, is interesting. If it implies that the major goal of urban life is environmental efficiency, however, or suggests that a city like New York or Beijing is simply a forest in drag, it is actively misleading. Cities, and human institutions generally, are far more complex than ecological analogs (primarily because of reflexivity and the overlay of cultural and technological evolution on biological evolution). The trick to using metaphors and analogies well is to understand not just what they say about similarities, but what they point out about important differences as well.

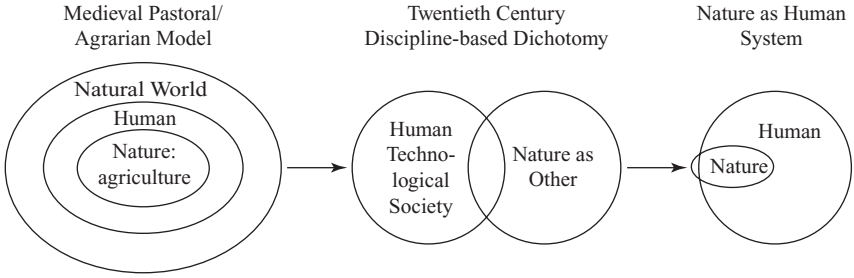


FIGURE 8.8 Evolution of Human/Nature Mental Models. This figure illustrates the progression from a world in which humans function within the constraints of nature, towards a world where nature is a design space and a commodity. In part, this illustrates one fundamental effect of technology: to render more and more of the "external"—nature and the environment, in this case—as internal to human culture, and technology. Some ideologies, such as environmentalism, are strongly committed to the technology/nature as other dichotomy, and strongly oppose any suggestion that nature is, in fact, increasingly a human domain (this is one basis for the European resistance to GMO agricultural crops). The initial question for the sustainable engineer is not whether such opposition is "right" or "wrong"—those are normative issues where the engineer's opinion is no less, and no more, valid objectively than someone else's—but whether it is realistic given the trends in technology, economics, and global culture today.

to being a human construct that we increasingly design and that is increasingly affected by the dynamics of human systems, especially technology and culture. Figure 8.8 illustrates this shift schematically: In a pastoral environment, it is apparent that natural systems do not do what humans want, and that human activity must fit into nature. If you were an early farmer and your crops failed or the weather didn't cooperate, you went hungry or died. Diseases were a constant threat to you, your animals, and your crops. Today in developed countries it is the opposite that is true: We are increasingly designing nature through, for example, building the landscapes we want and planting them with GMOs. If we don't recognize this fundamental shift in relationship—and we generally try not to—it is partially because our disciplinary structures and mental models are such that they actively discourage such a realization.

Figure 8.9 illustrates two mechanisms by which "natural" systems have become converted into primarily "human" systems. The first is the commoditization of landscapes, a process that began ten thousand years ago with the beginning of agriculture, but that accelerated dramatically with the rise of the modern state, the ability to own property, and the demand for land generated by the population explosion of the past two centuries. Other impacts of agriculture, such as changing hydrologic regimes to support production of food and fiber, have reinforced the "humanization" of natural

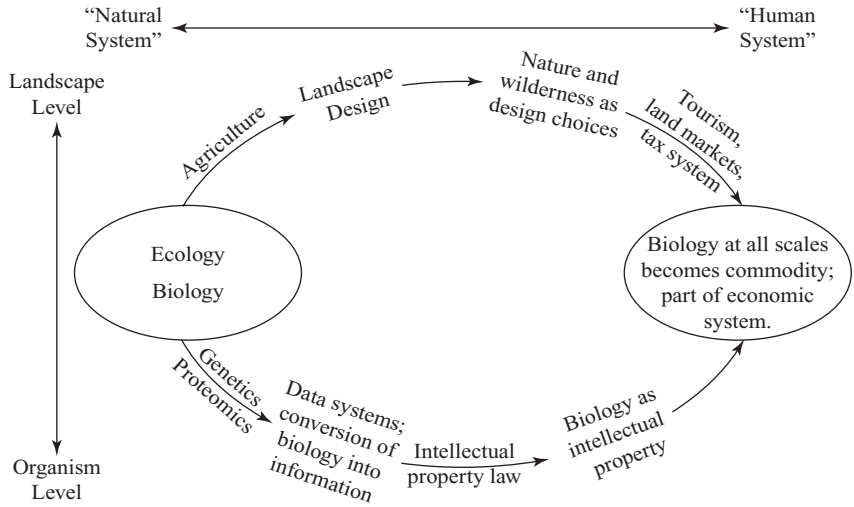


FIGURE 8.9 Evolution of Biology from "Natural" to "Human" System. This figure illustrates two paths through which "natural" systems, in this case biological systems, are being converted into "human" systems. The boundary between these two types of systems is fuzzy and increasingly arbitrary, which in some ways is indicative of how far along the process is. Similar paths can be constructed for many other "natural" systems: the hydrologic cycle, the nitrogen, phosphorous, carbon cycles, and the like.

systems. The second is the conversion of biological organisms into genetic and proteomic information, which is then converted into intellectual property, which then becomes a commodity and part of the economic system. The role of ICT, both as enabling technology and as the product into which wetware (biological systems) is converted as it is commoditized, is explicit in the latter system and implicit in the former (e.g., land cannot become owned until it is defined through some sort of information system, such as property law and records, frequently requiring the development of appropriate institutions, such as taxing authority and a state able to enforce ownership of land).

The commoditization of nature into human systems, often through ICT mechanisms and services, is perhaps one of the most basic and least appreciated actions of technology on the environment. But there are even more subtle questions involving ICT that a sustainable engineer might want to consider. For example, there are some technologies that are psychologically and socially strongly associated with freedom; such technologies are highly resistant to regulatory or environmental challenge for this reason. Currently, the technology that most people associate implicitly with their sense of personal freedom is the automobile; it is without question the predominant icon

of individual freedom in the world (look at automobile advertising from this perspective). This is a major reason why regulating things like vehicle gas mileage in the United States or speed limits on the autobahns in Germany is so difficult.

Consider, however, the possibility that several of the global trends we have already discussed could come together to fundamentally change this association between freedom and automotive technology. Begin with the enhanced functionality of ICT systems, especially the social networks and underlying technologies discussed above; add in accelerating access to bandwidth; and mix in demographic shifts such as urbanization (which make automobiles less functional because of congestion and close availability of stores and other resources) and aging populations (it is easier and safer to make ICT, rather than personal automobiles, available to many seniors). The result is the possibility that ICT rather than automotive technology could become the technological system associated with personal freedom. Were personal automobiles to become less psychologically important to individuals as a result, could it facilitate management of environmental impacts arising from transportation? Would it make, for example, speed limits on the autobahn more politically feasible? Would it encourage consumption of electrons or photons rather than fossil fuels? Might it be possible to reposition the automobile as primarily a source of transportation rather than an environmentally problematic emblem of personal wealth, power, hormonal status, and the like? (We have much more cultural freedom to seek economic and environmental efficiency in the first case, rather than the latter.)

Such thoughts reinforce several basic points for the sustainable engineer. First, in almost every sphere, the connection with ICT is becoming more important from a number of perspectives. From the positive perspective, ICT can clearly generate environmental and social efficiencies, but this comes with costs that must be understood and minimized—especially potential costs in terms of increased systems complexity and unpredictability, and security. (The more an ICT network is connected to outside sources, the more useful it tends to be, and the more dangerous and easier hacking or hijacking is.) Second, ICT is conceptually difficult, particularly because it extends across both technical and sociocultural domains—a mapping that increasingly characterizes the anthropogenic world and must be understood by anybody who purports to be doing sustainable engineering. Finally, of course, these examples reinforce the crucial point that, despite the denials and challenges of fundamentalists and environmentalists, the anthropogenic world is a technologically rich world; managing it requires intelligent engagement with technology, not wishful thinking or denial. It requires, in other words, sustainable engineering.

One key area of scientific and technological advance that plays into many of the emerging technology systems already mentioned is cognitive science, or "cogsci". Many of the advances in robotics, for example, reflect increasing understanding of how nervous systems and brains are structured and behave. Technologies such as the "telepathic helmet"—a device already under development that would read electrical patterns in one brain, code them for communication to another device, which would either turn it back into electrical patterns directly or convert it to words—depend heavily on advances in cognitive science. This is a tricky area, however, for any sort of prediction; the experience of artificial intelligence (AI), where dramatic advances were expected early in the field's history, but were never achieved because cognition turned out to be much more complex than first anticipated, is an important caution. Moreover, there are difficult definitional issues: How should one characterize, for example, the way that Google so dramatically extends human memory, creating a cognitive system that includes not just the human elements, but vast swaths of Net technology, or the way that many militaries are building "augmented cognition" technologies (known as "augcog"), where technologies intended to scan the battlefield for threats are integrated into each soldier's cognitive systems? Military augcog has its civilian mirror in the technologies that car companies are developing to enable seniors to drive even if their cognitive functionality is no longer what it was by building cognitive function into the automobile itself. These developments, and many more, clearly illustrate significant progress in cognitive science (as well as perhaps indicating some pretty substantial limitations on current understanding), but they are difficult to quantify in the same way as, for example, biotech and ICT capabilities.

In this regard, Figure 8.10 makes several points. First, it illustrates that, to the extent replication of processing functionality is analogous to (or the basis of) cognition, the progress in ICT can be directly tied to the potential for electronic cognition in the near future. The announcement by IBM that they had designed a virtual cat brain to sit on their computer, which was disputed by some on the grounds that we didn't yet know how to "duplicate" a cat brain, was in an important sense an indication that the physical capability to support virtual cognitive systems is rapidly being developed. In fact, given that microprocessor power lags supercomputer power only by some ten years, as Figure 8.10 demonstrates, perhaps the most curious possibility is that we will all be walking around with cat brains, in the form of our personal digital assistants (or "iCats"?), by somewhere around 2015, about the same time that supercomputer performance may enable the virtualization of a human brain (if the state of cognitive science can support such a gesture by that time).

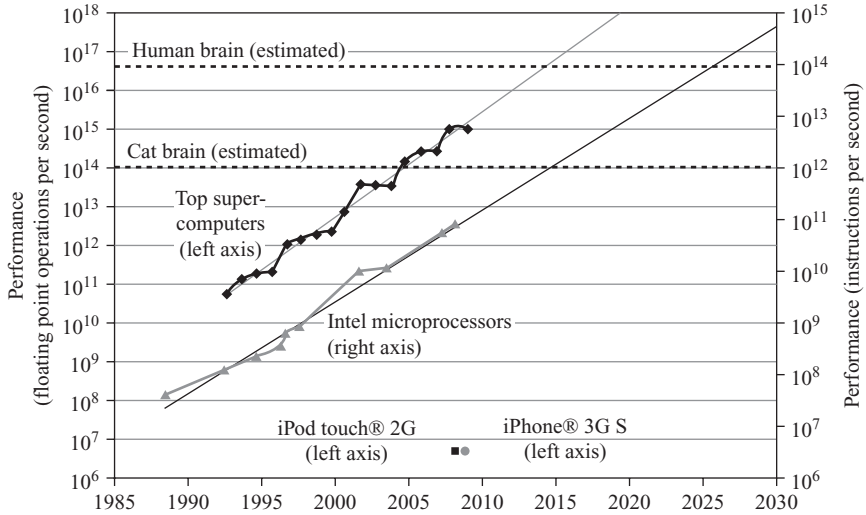


FIGURE 8.10 Computing Performance by Release Date. This graph illustrates the relationship among supercomputing power, microprocessor power, and human and cat brain functionality. It should not be read to imply that we understand how to virtualize the human brain, or that critical cognitive functions, such as consciousness, are understood; merely that the capability to match what we think the performance parameters of the human brain are, could be developed by 2025. This figure implies that chips with the power of a cat brain could appear in a decade and chips with the power of a human brain within a decade after that, although any such projections should be regarded as scenarios not predictions. Note that human and cat brain estimates apply to both the left and right y-axes.

(Data from: Ananthanarayanan, R., S. K. Esser, H. D. Simon, and D. S. Modha. "The Cat is Out of the Bag: Cortical Simulations with 109 Neurons, 1013 Synapses." Portland, OR: ACM, 2009; EconomicExpert.com. "List of Intel Microprocessors." http://www.economicexpert.com/a/List:of:Intel:microprocessors.htm: accessed December 17, 2009; Moravec, H. "When Will Computer Hardware Match the Human Brain?" http://www.transhumanist.com/volume1/moravec.htm: accessed December 17, 2009; TOP500.Org. "TOP500 List Releases." http://www.top500.org/lists: accessed December 13, 2009.)

8.3 THE HUMAN AS DESIGN SPACE

It will be immediately apparent from the above discussion that a major effect of the Five Horsemen is the profound change such technological evolution implies for the human (see Box 8.1). These changes cannot be easily categorized, either: Some human design technologies involve pharmaceutical enhancements of various cognitive functions; some involve exoskeletons that extend human strength and agility; some involve integration of human and technological cognition (e.g., Google) in ways that are not currently well defined, or even perceived; some involve redesign of human genomes. And there is always the opposite, because the general rule that anything that can be enhanced can be de-enhanced must always be remembered—thus design of

weapon systems that rely on de-enhancement of opponents in some way. The underlying change in perspective, however, is noteworthy and should be explicitly understood: The human being shifts from something that can be assumed to be stable and unchanging, to something that is subject to design and improvement, the same as any other physical system. This is a dramatic shift with significant implications across many domains.

It dramatically increases the complexity of the evolving world, for instance, because it is not just policies, institutions, and social and cultural frameworks that now can be deliberately changed as new information and knowledge is developed, but humans as well. Reflexivity—the incorporation of new information into the system from which it was created, thereby changing that system—has been a process that sociologists have dealt with at the level of communities and social systems. It now becomes inherent in human biology. Design changes in the human create new systems conditions that then drive new design changes in the human.

Moreover, there is a tendency in the literature on human enhancement, and indeed in the underlying research, to consider the human individual apart from his or her social and cultural context. Thus, for example, the implications of cognitive enhancement pharmaceuticals are evaluated based on their effect on individual performance and long-term neurochemistry. This is no longer adequate. Redesign of the human is not just an individual choice, because it does not just affect the individual. If one student takes cognitive enhancers to increase performance on critical professional tests, it may not change social dynamics; if many do, there comes a point where not to enhance ensures inadequate performance compared to the norm. Human redesign changes not just human individuals, but society and culture as well.

Lest the importance of this be lost, consider the basic point that the evolution of human technological competency such that the human itself becomes a design space clearly challenges cultural assumptions about appropriate boundaries between the sacred and the human for many belief systems. Such an impact can clearly be seen with the suite of human enhancements that, taken together, suggest the possibility of radical life extension. At one extreme, some predict the achievement of "functional human immortality" within fifty years, either as a result of continuing advances in biotechnology, or as ICT and computational power enable downloading of human consciousness into information networks. Although such predictions are viewed by many experts as highly unlikely, there is a growing consensus that substantial extensions of average life spans, with a high quality of life, are achievable in the next few decades. For example, the IEEE *Spectrum,* a mainstream technical journal, ran a series of articles in 2004 on engineering and aging, which concluded that using engineering methods and technologies to create conditions of minimal aging would allow average ages of well over 100, with a high quality of

life, within a few decades. What is left unsaid is the understanding that, given that sort of time frame and the acceleration clearly apparent in underlying Five Horsemen technologies, once one obtains a life span of over a hundred, it becomes increasingly possible that indefinite extension can be engineered—one just has to engineer people to live long enough for the next set of radical life extension technologies to be developed, in an ongoing process. Remembering that this possibility should be understood as hypothetical rather than a prediction, it nonetheless is apparent that the implications of such a development would be substantial. What happens to economic systems predicated on traditional paths of being born, being educated, working, retiring, and dying? What are the implications for relationships between generations—will thinking, for example, become ossified as older people, less integrated with the new technology of their time, stifle innovation by remaining in place too long and refusing to give up economic and political power to the young? And what of demographic policy? Birth rates and death rates are related; change one substantially, and it is necessary to reconsider the other unless indefinite population expansion can be managed. And, of course, there are deep religious questions: If humans become capable of providing functional immortality for themselves, how does the role of one's deity and one's religion change?

It is also interesting that, even though scientists and technologists are actively working on radical age extension, those in other areas of science, in policy, and in the environmental and sustainability communities remain generally unaware of these possibilities, despite their obviously challenging implications. This points out one of the important roles of sustainable engineering: As such possibilities pass from being hypothetical, to perhaps possible, to increasingly probable, social and institutional systems need to be kept informed so that they can adjust to the extent possible. Sustainable engineers need to be leaders not because it would be nice, but because it is increasingly critical: the evolution of human systems in a rapidly changing world where humans are themselves design spaces otherwise may become catastrophically unstable.

It may be that the most obvious response to this sort of development is to demand a halt to human enhancement; indeed, some conservative religious figures and a number of environmentalists have done so. But this is problematic for several reasons. First, it generally represents an attempt to impose a particular belief system on everyone else by restricting their access to potential benefits, including in some cases medical benefits (consider the stem cell controversy in the United States). Not everyone will accept such a tradeoff, especially if they are not believers in the worldview being imposed. Second, it is not at all clear that such a demand will work. Enhancing humans is likely to have such a powerful effect on relative cultural power that it is highly unlikely that all cultures would reject such technologies. If, for example, a nation could improve the average cognitive functioning of its population by only a little bit

on average, it would significantly enhance its competitive advantage over others. (This can be seen, albeit in primitive form, in the way countries jostle to improve their educational systems.) In addition, human enhancements are popular, even when they may come with price tags. Many athletes routinely use illegal and somewhat risky drugs to improve their performance; many students use off-label cognitive enhancers; and cosmetic surgery with little immediate health benefit, such as breast augmentation, is highly popular. People like enhancing, which is why those who argue against it generally do not attempt to change individual behavior, but try to convince governments to prevent research on or access to enhancements through coercion. And it cannot be forgotten that enhancements may indeed be preferable: One generation's enhancement is the next generation's medical technology, and a healthier life with less pain is not necessarily a bad thing. Finally, many human enhancements develop in military situations (see Chapter 11); this means that it is extremely hard to stop them. The argument that one should accept greater battlefield casualties rather than enhance war fighters is a difficult one; it becomes more so because many of the more powerful technologies may well be treated as national security secrets, and thus developed and introduced before any challenge is possible.

This does not mean that human enhancement is not a complicated and challenging area ethically and operationally. Many enhancements carry both risk and benefit, and undue optimism and pessimism may impose unintentional costs as a result. Furthermore, it is true that on the eve of humans becoming self-reflexive design spaces we don't understand nearly as much about humans as we would like to. There do appear to be times when introspection, even suffering, leads to a deepening of the human experience and a broader understanding of the human condition. The challenge is to understand how humans function without necessarily glorifying suffering or pain for its own sake, an especially difficult task as such psychological frameworks are, of course, at some point part of the design project as well. That we "need" suffering now to become fully human or to reach creative heights does not logically imply that a redesigned human would require suffering in the same manner.

One of the key insights of sustainable engineering is that technology is not just an objective engineering domain, but has substantial social, cultural, and psychological dimensions as well. Thus, intercultural differences can be important, especially with emotionally charged technologies such as human enhancement. For example, human enhancement is often viewed, particularly by opponents, as some sort of victory of technology over the human, as if each were a separate domain. The Enlightenment Romantics had their Frankenstein narrative, and this remains powerful, particularly in Europe and, more generally, in the West. On the other hand, some Asian cultures

seem to have an easier time accepting continued integration of the human and the technological: in Japanese anime movies, for example, it is usually not technology per se, but the evil people behind technology, who are identified as problematic. And, indeed, work by philosophers such as Clark suggests that the current opposition to human enhancement is at best a temporary opposition; what we can anticipate is not that the human and the technological will clash, and one will emerge victorious, but rather that the two will continue their long process of merging. This doesn't mean profound changes won't occur, especially in older concepts of what constitutes "the human," or that such change won't ignite serious cultural clashes. Nor does it mean that, especially if the human family continues to fragment into "varietals"—as, indeed, the "digital natives" who are comfortably embedded in their ICT networks may already be—social and institutional conflict may not at times be intense. But it does mean that those who choose to engineer in a world increasingly defined by accelerating technological evolution need to be more sophisticated and more versed in the understanding and management of technology as a profoundly human activity, one that now is changing not just natural systems and the built environment, but the human itself, in complex and unpredictable ways.

Exercises

8.1 What characteristics of nanotechnology are likely to pose concerns for sustainability? What are the best possible ways to mitigate these concerns (a) in the laboratory; and (b) in the field?

8.2 You are working on a chip that, if implanted in the brain of a paralyzed individual, would allow that person to interact with a suitably prepared suite of technologies, thereby mitigating the effect of the paralysis. Your funding agency, however, might use this technology to enable remote control of major weapon platforms. What sorts of ethical questions does this raise for you?

8.3 Advances in biotechnology and cybernetics are fast closing in on the production of sentient beings that can replicate many human characteristics. Similarly, humans will soon have "machine" parts that may serve as advanced prosthetics that give them new abilities and senses. What does this mean for the term "sustainable engineering"?

8.4 Science fiction novels and movies tend to have a dark, despairing side. Why do you think this is so? What motivates writers to express a sense of loss, rather than excitement about the possibilities of the future?

Annotated Bibliography

1. There are a number of books that provide an introduction to the Five Horsemen, although none use that name. *Radical Evolution* by Joel Garreau (New York: Doubleday, 2005) provides a readable and balanced introduction. *The Singularity is Near* by Ray Kurzweil (New York: Viking, 2005) is more optimistic in its suggestion that these technologies will rapidly lead to a "singularity," a point beyond which most of what we know is obsolete. The volume does, however, marshal a substantial amount of data and analysis in presenting this thesis that is worth understanding even if one disagrees with Kurzweil.
2. Andy Clark makes the important point that humans have always been integrated wetware/technology systems in his book *Natural-Born Cyborgs* (New York: Oxford University Press, 2003).
3. Two relatively optimistic books on human enhancement include *Converging Technologies for Improving Human Performance* by M. C. Roco and W. S. Bainbridge (New York: Kluwer Academic Publishers, 2003), and *Citizen Cyborg* by James Hughes (Boulder, CO: Westview Press, 2004). Although both make persuasive reading, it is worth asking as you go through them whether humans really are as simple and as independent from community as such approaches may assume.
4. Although it isn't a big part of the discussion in this chapter, it is worth thinking about, for those with such a bent, what exactly is the core of "self" and "human" actually. What, for example, changes as memory is offloaded onto the Net thanks to powerful search engines, or we increasingly implement augmented cognition not just in combat zones, but, for example, as assist systems for older populations that continue to use automobiles (a set of technologies that most major automobile firms are, in fact, working on right now). Some provocative ideas on such topics can be picked up from *Godel, Escher, Bach* by Douglas Hofstadter (New York: Vintage, 1980), *The Mind's I* by Douglas Hofstadter and Daniel Dennett (New York: Bantam, 1981), and *Freedom Evolves* by Daniel Dennett (Viking, 2003).

CHAPTER

9 Green Chemistry

9.1 INTRODUCTION

The first and most obvious point about "green chemistry" is that it is self-evidently not **sustainable chemistry** because it focuses on the environmental domain rather than the broader sustainability domain. But green chemistry remains an important focus for environmentalism and consequently for sustainability advocates, which can be attributed to historical and ideological reasons. For one thing, concerns about the environmental effects of the chemical industry were an initial spark for the environmental movement in the United States and elsewhere—after all, Love Canal, a community in New York where a school was built on top of a chemical waste site, was the incident behind the establishment of the Superfund waste site remediation program (technically known as the Comprehensive Environmental Response, Compensation, and Liability Act, or CERCLA, passed in December 1980, by the U.S. Congress). A few well-publicized and major failures of chemical facilities, such as occurred in Bhopal in India, have added to such concerns. For another thing, chemicals to many people are somewhat frightening because they are arcane and unfamiliar, with names that make little sense and sound vaguely technocratic. In addition, just as in the case of industrial ecology and other existing practices and methods, the evolution from environmental concern to sustainability is a work in progress that is not yet complete. Accordingly, this chapter will explore green chemistry, with the last section expanding that discussion to cover some observations regarding sustainable engineering.

To begin with, especially from a chemical engineering perspective, green chemistry and the application of industrial ecology principles to chemical engineering is not new. Indeed, much of the history of the chemical industry consists of learning to turn the "waste" from existing processes into feedstocks for new products, and managing the environmental impacts of waste streams until that happens (often under direct or indirect pressure from government regulations). Consider, for example, the production of soda (sodium carbonate), needed for glassmaking, dyes, soapmaking, bleach, and other purposes. This was an important process early in the Industrial Revolution. Nicolas Leblanc developed a method for making soda using calcium carbonate, coal, and sodium sulfate that he patented in 1791. The Leblanc process, however, produced waste streams with serious environmental impacts, including hydrochloric acid fumes and calcium sulfide. Failure of industry to properly manage this waste stream led to regulations, such as the 1863 Alkali Act in England, which required English soda producers to recover most of the hydrochloric acid.

But waste streams are in one sense simply potential resources that we haven't been smart enough to use yet, and it wasn't long before chemical engineers learned to convert the hydrochloric acid to chlorine by catalyzed air oxidation. This chlorine could then be reacted with lime to produce bleaching powder; the availability of cheap bleaching powder then led to a process developed in the 1880s that used bleaching powder to process North African esparto grass for paper. Meanwhile, however, the original Leblanc process was being replaced by a more efficient ammonia-soda process developed by Ernest Solvay, which was in turn displaced by the development of electrolytic soda production in the late 1890s. Such patterns of increasing efficiency and repurposing of residual flows into product are not just common to, but perhaps the core of, the evolution of the modern chemical production sector.

The sustainable engineer should in passing note two very important characteristics of chemical manufacturing that makes such a socially and environmentally preferable evolutionary process possible. First, because chemical synthesis systems are generally under the control of a single firm or group of firms, whatever efficiency and cost benefits accrue from redesign of operations and reuse of residuals flow to the firm, thus providing a direct incentive to continue such innovation. If, instead, the benefits went to a recycler or another entity, those incentives would be substantially smaller, or nonexistent. Moreover, if a residual flow is under control of the same firm that intends to reuse it, quality control is much easier; if a firm is using a residual flow from a third party, quality can often be a problem. Few firms, after all, carefully design their waste flows to maintain their

quality. Second, the costs to a chemical firm of a mistake or incident—the classic example being Bhopal—are huge, and under existing legal regimes, firms can often be sued not just in the country where the incident happened, but in many others as well. (Union Carbide was sued for the Bhopal disaster, which occurred in India, in American courts as well as Indian courts.) This does not assure operations are always safe, of course, but it provides strong incentives both at the firm level and, because the public often does not differentiate between companies in a particular sector, at the sector level as well.

So if green chemistry is not new, why does it need to be invented, especially if the concept does not represent a dramatic break with traditional chemistry or chemical engineering practices? One way of understanding the green chemistry movement is simply as a shift in emphasis. In a world requiring ever more chemistry and chemical industry activities because of population and economic growth, a more formal emphasis on environmental impacts is seen by many as necessary. From this perspective, green chemistry is not a long-range research and operational field, but rather a transition, lasting until traditional chemical and materials science programs and practices internalize the potential of new, more environmentally efficient technologies such as enzyme-activated biotech production processes. At that point, perhaps like sustainable engineering itself, green chemistry simply becomes "good chemistry."

But this perspective hides a more provocative possibility: green chemistry and its agenda of integrating a more "environmental" or "ecological" approach may be an indicator of a fundamental shift away from the historical reductionist emphasis of chemistry and material sciences, and indeed of scientific disciplines as a whole. In this view, green chemistry might be seen as a small step towards augmenting traditional approaches with a much broader earth systems engineering and management, and sustainable engineering, perspective. In doing so, chemical engineering would move from a fairly strong focus on immediate production and use of chemical systems to a broader understanding of the implications of chemical engineering on complex regional and global integrated human/natural/built systems, rather than specific elements of such systems. If this is the case, green chemistry itself has a long way to go, as much of its current focus remains firmly on bench scale considerations rather than systemic perspectives as Box 9.1, which discusses the U. S. EPA principles of green chemistry, illustrates.

Nonetheless, it is probably no coincidence that, just as human activity scale to a point where many major material flows are increasingly impacted, the ideas and interest in green chemistry bloom. Consider, for example, the increase in production in commodity plastics that have occurred over the

BOX 9.1 Principles of Green Chemistry

Drawing from the literature, the U.S. EPA identifies twelve principles of green chemistry (see http://www.epa.gov/gcc/pubs/principles.html), which range from general admonitions to reduce waste, to more specific guidance on chemical engineering design. Paraphrased and augmented, these principles are:

1. Prevent waste by proper design of chemical syntheses.
2. Design chemical products to be fully effective, yet have little or no toxicity, thus inherently making them safer.
3. Design chemical syntheses and manufacturing operations to be safer, using less hazardous materials and processes where possible.
4. Use renewable materials where possible, and where using renewable makes sense otherwise (e.g., watch out for policies such as the U.S. corn-based ethanol scale up).
5. Use catalysts rather than stoichiometric reagents. Catalysts can be reused many times, whereas stoichiometric reagents are changed by being used, and thus a constant feed is required and wastes are usually generated.
6. Avoid using chemical derivatives (blocking or protecting groups, which must subsequently be stripped, usually creating waste). This is a specific instance of a more general rule: Use the simplest chemical pathway, all else equal.
7. Use the chemical pathway that puts the maximum amount of chemicals into the final product, all else equal. A more general rule is to maximize production by volume given inputs by volume; this may mean developing uses for residuals that might otherwise go to waste, and thinking of chemical synthesis as producing a palate of products, rather than single outputs.
8. Use safer solvents, which in many cases will mean catalytic or biologically based production systems.
9. Strive for energy efficiency in production, all else equal.
10. Design chemicals to break down after use. Obviously, this requires knowing what all uses of a particular chemical might be; one would not want a bioplastic put in a jet airplane where it happily begins to biodegrade before the jet does, but a bioplastic in a shopping bag might be perfectly appropriate. It also requires knowing something about the breakdown pathways, so one can be sure one is not causing more problems than one is solving by, for example, creating breakdown products that are more toxic than the original material (plastic bags may be unsightly, but they're not toxic).
11. Make chemical synthesis and production more information intensive, and thus more efficient systemically. This is a specific application of the general sustainable engineering principle that more real-time data enables more efficient operations of most systems and carries with it not just the advantages of enhanced efficiency, but also the cautions—one wouldn't want the production information network of a large chemical plant hacked, for example, so security is an important consideration here.
12. Design synthetic pathways and facility operations to minimize the chances of accidents, hazardous exposures to workers, and releases to the environment,

all else equal. This is a specific application of the general sustainable engineering principle that it is better to design technologies that are inherently failsafe than to rely on administrate controls.

These are useful guidelines, but it is important that the sustainable engineer understand their very significant limitations. To start, an underlying assumption in several of them—indeed, in many similar lists of design principles based essentially on heuristics, or "rules of thumb"—is "all else equal." But, of course, it is very seldom that changes can be made to one dimension of a complex adaptive system such as a chemical manufacturing facility without affecting much else, because of the coupled nature of the subsystems involved. Reducing energy consumption by lowering the temperature of a reaction vessel, for example, might lower the efficiency of the desired reaction, leading to less product and more waste. The essence of design is, after all, difficult tradeoffs, and the seeming simplicity of these heuristics must be taken with a large grain of sodium chloride.

Heuristics such as these also often rely on assumptions about what is environmentally preferable. Thus, for example, Principle 4 calls for the use of "renewable materials." Although this may often be desirable, the idea that "natural" or "renewable" materials are preferable is, at best, an over-simplification that must be tested by local conditions. Timber is renewable, for example, but the extensive logging in many areas of the world indicates that current levels of demand can often result in serious impacts on environmental systems and indigenous populations, regardless of whether a renewable is involved. Some chemicals, such as terpenes used in aqueous piece-part cleaning systems, are "renewable" and "natural," but when used at far higher concentrations than found in oranges and other citrus, can have significant health and environmental effects. Be critical of any blanket injunction regarding materials, even if it "sounds right" or is popular with the public, or even if it is recommended by U.S. EPA. Your job as a professional is to make the most sustainable material choice in very complex design environments; chemistry is not a popularity contest but a science.

Finally, note that these principles all implicitly assume that chemical engineering is bench scale. This is an important scale to focus on, of course, but it is also entirely inadequate if one is truly interested in green chemistry. The most obvious counterexample is chlorofluorocarbons (CFCs), which, according to these criteria, are the basis of the perfect green chemistry. Because they are chemically more stable, they are safer, usually use less energy, and cause much less damage to local environments than alternatives. Of course, at a systems level they deplete stratospheric ozone, which increases penetration of high-energy UV radiation from the sun to the Earth's surface, and damages DNA and organisms. This is why CFCs are now subject to phase out under the Montreal Protocol and enabling agreements. The point is not that sustainable engineers should not do better at bench-scale chemical engineering design, nor that the principles given above aren't useful. But it is important to understand that sustainable engineering requires going beyond the incremental and partial to systemic impacts and implications, across both environmental and human systems, in ways that are far more difficult and comprehensive than suggested by these principles. They are a good first step towards sustainable chemical engineering, but only that, and the sustainable engineer should expect to be far more sophisticated than they imply.

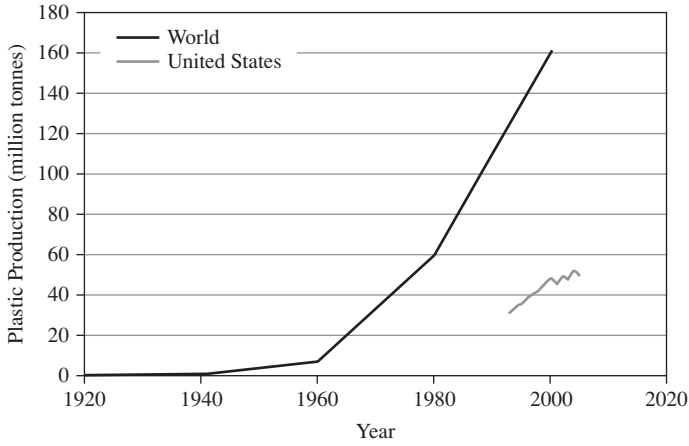


FIGURE 9.1 Annual Plastic Production. The exponential increase in production since the end of World War II is particularly striking and important because it is based on fossil fuel inputs and indicative of the expansion in material use, and thus integration of human, natural, and built systems, in the past decades. In historical terms, this is a remarkable change in a very short period of time.

(Data from: Commodity Research Bureau. *CRB Commodity Yearbook*. Hoboken, NJ: John Wiley & Sons, 2007; Elias, H. G. *An Introduction to Plastics*. 2nd ed. Weinheim, Germany: Wiley—VCH, 2003.)

past decades, as illustrated in Figure 9.1. Similarly, it is extraordinary how much of many primary materials are mobilized by humans; as was illustrated in Table 1.5, humans mobilize over half of the annual flows of materials ranging from the toxic (arsenic and lead) to the familiar (copper and iron). Equally of interest, even for major earth elements such as calcium, carbon, and chlorine, human mobilization is significant, even if small in relation to overall cycle size.

Green chemistry as a reaction to the Anthropocene and its overtones of human impacts on natural systems is thus understandable, but requires augmentation from the perspective of sustainable engineering. First, it is necessary to include the social dimensions of chemical engineering, which are generally regarded as highly positive, rather than simply emphasizing the negative impacts of chemicals and chemical engineering on the environment. Few would choose to live without modern drugs, food preservation, clothing, and material structures, so these are part of a sustainability analysis, even if they are not part of an environmental one. Second, the sustainable engineer must not downgrade the importance of scale or boundary selection, a point made more clear by the human experience with CFCs, once widely used in industrial processes for cleaning, fire protection, and refrigeration purposes. Even if some chemicals are useful in benchscale applications, as CFCs certainly were, they may present serious systemic risks as they are adopted across a globalized

economy that tends to rapidly adopt useful technologies at a large scale. Some of these points are further illustrated in Box 9.2, summarizing the recent history of arsenic use in the United States; another example that has already been discussed is corn-based ethanol as a substitute for fossil fuels, a chemical technology that is not inherently problematic, but becomes so when scaled up to commercial levels.

BOX 9.2 Arsenic Management in the United States

Arsenic is a toxic element found in many forms of varying toxicity in nature. (For instance, arsenic trioxide is far more toxic than elemental arsenic.) In biological systems, arsenic compounds interact destructively with a number of different metabolic pathways; acute exposures usually result in death by organ failure, while chronic exposures, such as in drinking water, may result in cancer, open sores, and other injury. Although arsenic-based drugs have been used up until the early days of the twentieth century, arsenic has long been known to be poisonous; indeed, it was sometimes called "inheritance powder" in the Middle Ages in Europe for obvious reasons (apply to rich uncle, await inheritance). Arsenic is produced commercially as a by-product of copper, gold, and lead mining.

An example of bench-scale chemical engineering having unanticipated (albeit easily foreseeable) scale effects is provided by the history of arsenic use in the United States from 1960 onward, illustrated in Figure B9.2.1.

Arsenic, a by-product of copper production, was historically used in the United States as a pesticide (and in much lower quantities as a dessicant). EPA regulations in 1987 and 1993 essentially phased out this use. However, although there was a dip in use as a result of these initiatives, the level of U.S. arsenic use did not diminish over time, as use shifted to wood preservation, primarily as chromated copper arsenate (CCA). Approximately 48 billion board feet, or 113,000,000 cubic meters of wood, were pressure treated with CCA from 1985 to 1999. U.S. consumption of pressure-treated wood at that point accounted for 90 percent of worldwide arsenic demand and made the United States the world's largest consumer of arsenic. At the turn of the century, U.S. demand for arsenic was being met entirely through imports, primarily from China (which at that point was supplying about 85 percent of overall demand).

On the one hand, wood preservatives serve the important function of extending the life of wooden artifacts, especially buildings; without them, deforestation would be far worse than it is. On the other hand, arsenic was banned from agriculture in recognition of the risks involved in dispersive use of toxic materials: spreading arsenicals on fields where they can blow or wash away is a poor idea, just as putting lead in gasoline where it is ejected through the tailpipe is (these are good examples of a toxic material becoming hazardous, because use patterns generate exposure). In this light, wood treated with CCA is little improvement, because it

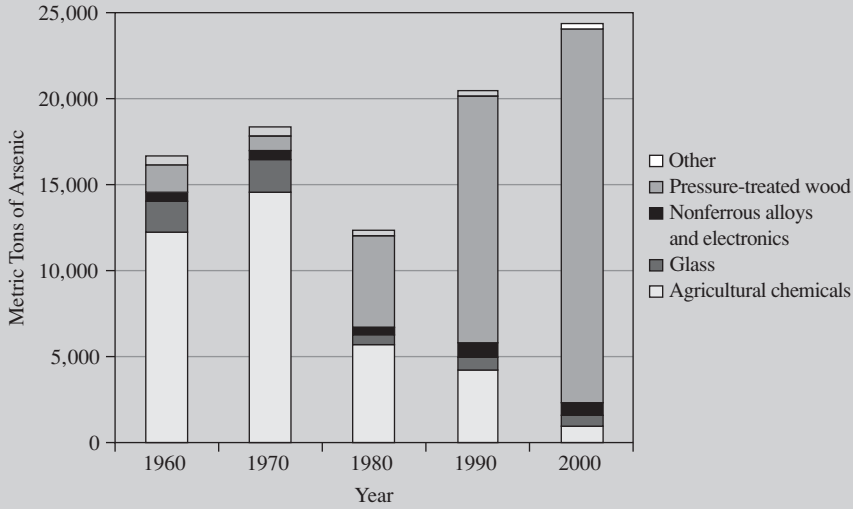


FIGURE B9.2.1 United States Arsenic Use. This figure illustrates the dramatic shift from agricultural use to use in pressure-treated lumber, to the point where overall arsenic consumption increases after 1980.

(Data from: Loebenstein, J. R. *The Materials Flow of Arsenic in the United States*. Washington, DC: United States Department of the Interior, 1994; Matos, G. R. and W. E. Brooks. "Historical Statistics for Mineral Commodities in the United States, Data Series 2005-140: Arsenic End-use Statistics." http://minerals.usgs.gov/ds/2005/140/#arsenic: accessed October 18, 2009.)

generates significant demand for arsenic, and uses it in a manner, wood preservation, that tends to disperse it across the environment. (Very little treated wood is recovered from old uses and recycled properly; most is disposed of in dumps not designed to hold hazardous materials or it simply degrades in the landscape.)

From the perspective of sustainable engineering, this case leads to two observations. First, while one may be rightly concerned about this use of arsenic, it is not necessarily the worst choice. To be able to make that determination, one would need to evaluate the social, environmental, and economic implications of the alternatives, from much higher demand for untreated wood (which would not last nearly as long as CCA-treated wood), to different possible treatments. Each one of these would no doubt have some negative environmental implications. (Preservatives, by and large, act by being toxic to organisms that would otherwise break down the material involved.) In short, there is a serious data-intensive question involved in the design choice.

Second, this case reaffirms the importance of understanding scale effects in a globalized economy. Chemical engineering at bench scale may provide many interesting possibilities, but with commodity markets such as lumber involved, it is important to remember that what may be a small and easily managed concern at bench scale can rapidly become a serious environmental or social concern once commercialization of the design is achieved.

9.2 THE CFC CASE STUDY

Green chemistry emphasizes the selection from the beginning of materials that are safe, are inert except for necessary chemical properties, enable reduced energy consumption, and are energy and material efficient. This, in a nutshell, describes CFCs, a class of halogens that were first synthesized by Thomas Midge of General Motors in 1928. CFCs were indeed designed to be substitutes for much more toxic and dangerous materials, including ammonia, methyl chloride, and sulfur dioxide, that were then being used for refrigerants and that had resulted in a number of fatal incidents in the early 1920s. They proved to be inert and safe materials in a number of applications, including not just for refrigeration, but also for industrial and consumer cleaning systems, as propellants for aerosol sprays, for inert atmospheres for fire protection in sensitive areas, and the like. They were also relatively inexpensive. For all these reasons CFCs were thus widely and quickly adopted, and perhaps ironically, could have been the poster child for green chemistry had such a concept existed in the decades after the 1920s. Figure 9.2 illustrates some of the more popular CFCs: an "R" designation means the chemical was primarily used as a refrigerant.

As with many technologies, the varieties and scale of uses of CFCs became apparent only after their introduction; in general, this chemical family became exceedingly popular because of its safety and effectiveness in many applications. In large part, these characteristics arose from the notable

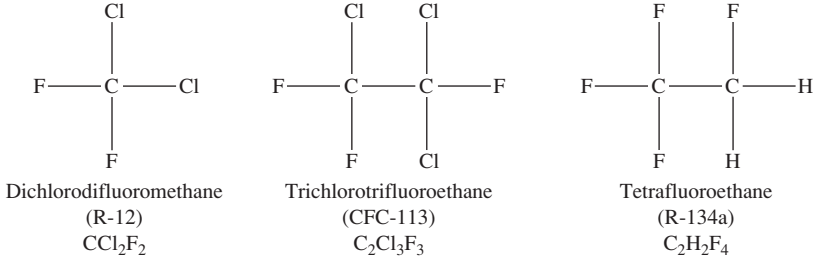


FIGURE 9.2 Chlorofluorocarbon Molecules. This figure illustrates some of the more popular members of the CFC family. CFCs are named according to a fairly arcane system developed by Thomas Midgley and A. Henne, where the generic identifier, CFC-01234a, reflects the chemical makeup of the particular compound. Starting from the left, the first numeral represents the number of double bonds and is omitted if there are none; the second gives the number of carbon atoms minus 1 and is omitted if that number is zero; the third stands for the number of hydrogen atoms plus one; the fourth gives the number of fluorine atoms and the fifth the number of chlorine atoms replaced by bromine. The small letter at the end is used to identify isomers. Thus, R-12 can be broken down as 0 double bonds, 0 for carbon (1-1), 1 for hydrogen (0+1), 2 for fluorine, and 0 for the number of chlorines that have been substituted for by bromines. Don't worry about the formula, because most anthropogenic members of the family are being banned; it is of historical interest, but should not be on your final exam.

stability of these compounds under ambient conditions. Unfortunately, this stability lasted until the materials migrated to the upper atmosphere, where they underwent photolytic decomposition by UV radiation, freeing up chlorine radicals that catalyzed the destruction of stratospheric ozone. (Remember that a catalyst encourages a chemical reaction without being changed by it, so each chlorine radical could destroy many ozone molecules.) Because stratospheric ozone absorbs high-energy UV radiation that would otherwise reach the Earth's surface, causing substantial harm to biological systems, the need for reducing use and thus release of CFCs (and certain related compounds) became apparent. Accordingly, international negotiations led to appropriate legal mechanisms for doing so, in the form of the Montreal Protocol and its subsequent amendments. This process was supported by industrial producers of CFCs, industrial users of CFCs, many nations, environmentalists, and most NGOs involved in the process, and has thus been an outstanding success story in terms of earth systems engineering and management.

This is an extraordinarily rich case for the sustainable engineer. To begin with, it obviously highlights the shortcomings of green chemistry, in that perhaps the most "green chemistry" product of the twentieth century turned out not to be environmentally preferable at all. The reason for this is critical: The "green" properties of a chemical system cannot be known unless impacts at all scales are monitored and thus perceived. (The importance of scale, especially in a globalized economy, is seen in the case of arsenic use in the United States, discussed in Box 9.2.) This is extremely hard when the effect—in this case, the reaction of an industrial chemical with ozone in the stratosphere under very different conditions than those for which it was originally designed—is unpredicted and unexpected. Indeed, the two scientists that first suggested the possibility of this arcane chemistry in 1974, Mario Molina and F. Sherwood Rowland, received a Nobel Prize for their work in 1995. Even if in an excess of caution a monitoring system focused on CFCs had been set up when they were introduced, it is not clear that the stratosphere would have been an obvious domain to begin monitoring. This case clearly illustrates that sustainable engineering in the chemical sphere remains somewhat of an art form and must rely on active exploration of potential systemic effects that, according to received wisdom, may not be obvious or anticipated. It also means that NGOs, which frequently come up with negative scenarios as part of their activities, serve a useful purpose to the extent they develop and present data indicating problems or issues where none were anticipated (a function somewhat undermined when public relations campaigns such as Frankenfood supplant serious environmental activity). It also means that sustainable engineering must consistently return to fundamental systems, whether natural, human, or built, and explore unexpected perturbations and changes, to determine whether unanticipated

impacts of engineered systems are expressing themselves and, if so, what actions may be required in response.

Another important question from the sustainable engineering perspective is: What went right in the CFC case? Here, it is important to recognize that in many of the technologies where CFCs were used, engineers had explored a number of additional options, from aqueous cleaning systems using engineered detergents and substances such as terpenes, to more hazardous materials with less systemic effect, such as flammable propellants for spray cans. Thus, technological evolution since the introduction of CFCs had created a rich option space, so that alternatives for many CFC uses were already developed, even if they were on the shelf because they couldn't compete with the many advantages of CFCs. The lesson here is that robust technological option spaces are a critical and often overlooked resource for an engineer, a firm, or a society seeking flexibility to adapt to an uncertain technological future. Banning CFCs would have been far more difficult and contentious if there had been no available substitutes: This is not just a political reality, but also a matter of cost/benefit analysis. When options are available, it reduces the cost of shifting technologies; when they are not, and the choice is, for example, between using CFCs and halting production of military systems, the costs and benefits are quite different.

This also applies institutionally. In the case of CFCs, the same interests that manufactured and used CFCs were able to benefit from, or at least transition at low cost to, these alternatives. This rough confluence of costs and benefits was also helpful in facilitating the transition. Importantly, the major manufacturers of CFCs were also in a position to manufacture and thus benefit from many of the alternatives. So, rather than opposing the phase out of CFCs, these firms were in support of it. Industrial users, seeing that their operations would not be unduly affected, were also supportive. Accordingly, the usual configuration of "old technology" interests opposing new, replacement technologies did not occur in this case. (It is important to note that opposition to change at the early stages of dialog about issues such as ozone depletion need not be unprincipled: in most scientific discussions at this stage, it is usually not clear how serious and immediate the proposed perturbations are, and supporting data that could resolve the issue definitively are generally not available. In short, opposition to change, especially if it carries other costs such as safety implications for manufacturing employees and higher energy consumption, is not unreasonable or unethical, except in the usual heated rhetoric that accompanies such discussions.) Compare this institutional alignment with the widely scattered and powerful interests involved in climate change discussions, and you begin to get an idea why the CFC negotiations went smoothly and the Kyoto Protocol discussions haven't.

Another observation about this case is that it validates the need for sustainable engineering and for an earth systems engineering and management approach. Here, a small human activity, the use of a relatively minor class of

materials in volumetric terms, destabilized a major earth system, the stratospheric ozone layer, with potentially disastrous implications. (It is likely that human skin cancer rates increased as a direct result of the weakening of the protective stratospheric ozone layer, for example.) And note that the feedback from perturbation of the natural system rippled quickly and profoundly into human technological, political, and economic systems as well. Understanding such complex networks of effects rigorously is not something our essentially reductionist approach to chemistry, or research generally, easily facilitates; moreover, such effects are far beyond the remit of green chemistry as currently practiced (see Box 9.1); they cannot, however, be beyond the scope of the sustainable engineer's concerns.

In some ways, this case also supports a sustainable engineering approach; after all, the controls on CFCs have worked, and the potential for a catastrophic shift (from our perspective) in a major earth system appears to have receded somewhat. In other words, given adequate knowledge, society is not helpless, and these highly complex integrated human/natural systems can be designed and managed. But it must be admitted that there seems to have been an element of luck involved in detecting the stratospheric ozone depletion in time to respond rationally; in the future, it would be unwise in the extreme to depend on such luck, especially as more complex, integrated, and co-evolving systems become engaged (as in climate change). Moreover, green chemistry as currently framed is not enough. We will need to move towards a more comprehensive sustainable engineering approach, including developing the competency to:

1. Determine, on a real-time basis, what systems are relevant to particular chemical families.
2. Dialog with those relevant systems at all scales.
3. Prioritize values, ethics, and goals so that inevitable conflicts among them can be resolved and the resulting decisions accepted (this is the systems engineering aspect of sustainable engineering).
4. Develop robust technology option spaces for critical materials and technologies, so that if materials turn out to be unexpectedly problematic, changes can be made effectively and with minimal social cost and political opposition.
5. Support development of the institutional capacity to carry out these activities.

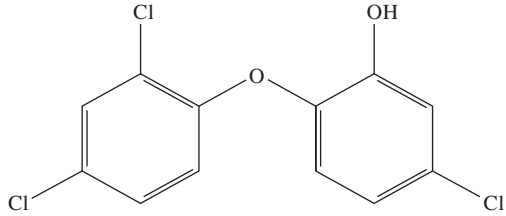
9.3 ANTIMICROBIALS AND PHARMACEUTICALS AS EARTH SYSTEMS

Another example of dysfunctional disciplinary and intellectual boundaries given real-world physical behaviors is provided by the flow of designed chemicals such as drugs or antimicrobials, or their metabolites, through the organism

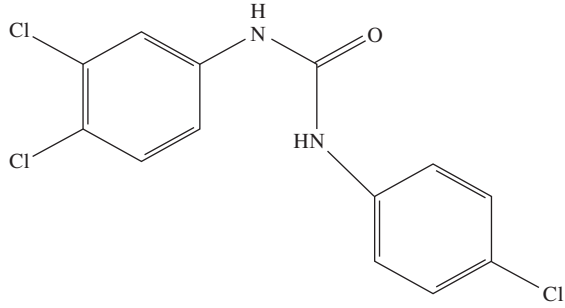
for which they are designed, and then into regional and global earth systems. The extent of this phenomenon is increasingly apparent as studies show that everything from cocaine and other recreational chemicals, to estrogen mimics, to large volume antimicrobial materials are not "consumed" in the human body, nor are they, or in some cases their metabolites, rendered biologically inert by water treatment systems (where robust water treatment facilities actually exist). Rather, they survive engineered treatment processes to flow into the natural environment, where they are widely dispersed through, for example, placement of treated waste on agricultural land, or simply through release back into surface and underground waster systems. At least in some cases, the release of manmade chemicals into the environment has the potential to disrupt natural systems.

Nonetheless, this growing phenomenon is for many reasons unaddressed. This is not a simple failure of science or regulation; rather, it reflects the broad challenge to institutions or society of changing fundamentally from the bench-scale mental model we discussed in regard to CFCs, to a different paradigm that understands that design of chemicals in a globalized economy is, in fact, design of earth systems at many scales. Failure to respond ethically and rationally to this phenomenon thus arises from many sources, including weak data and uncertain science that are constantly evolving, institutional capability, disciplinary boundaries, economic and financial incentives, and the inherent complexity involved in integrated human/natural/built earth systems. In addition, differing regulatory regimes are almost never integrated to address new conceptual challenges effectively or comprehensively. In the United States, for example, the federal Food and Drug Administration is responsible for manufacture and efficacy of drugs in humans, but the EPA is responsible for anthropogenic chemicals in the environment. The two seldom manage to communicate effectively internally, much less across institutional boundaries, especially where each has different but powerful interest groups to satisfy. NGOs, after all, intentionally reflect very specific issue orientations, and an NGO established to advocate regarding a particular disease is going to have very different interests than one concerned about the presence in the environment of metabolic byproducts of a drug treatment for that disease. Under these circumstances, it is not surprising that comprehensive social, economic, and environmental dimensions of designed chemical earth systems are poorly defined, and the appropriate balances among costs and benefits much less so.

Take as an example two popular antimicrobials widely used in consumer soaps, detergents, cosmetics, toothpaste, mouthwash, cutting boards, food packaging, carpets, and even toys: triclosan (TCS) and triclocarban (TCC) (shown in Figure 9.3). These materials are as pervasive as they sound, even if generally found in fairly low quantities in each application; it is almost impossible to find a home in the United States that does not routinely wash some



TCS
Triclosan, $C_{12}H_7Cl_3O_2$



TCC
Triclocarban, $C_{13}H_9Cl_3N_2O$

FIGURE 9.3 Triclosan and Triclocarban. This figure shows the chemical structures of triclosan and triclocarban, two antimicrobials common in many consumer products.

quantity of TCS and TCC down the drain every day. Moreover, like many drugs, hormones and other designed biologically active chemicals (or their metabolites, which in many cases remain biologically active), TCS and TCC are not removed by wastewater treatment systems. In fact, one nationwide survey in the United States indicated that almost 60% of sampled streams had detectable levels of TCS. Another U.S. study of major rivers in 13 states found that TCC was detectable in all samples collected downstream of wastewater treatment plants. Again, like many other designed materials, these chemicals remain active once released. Both of them persist in sediment, for example, and they continue to have significant toxic effects on aquatic microbial communities—which is not surprising given they were designed for just that. Moreover, it is not just direct release that is important; some studies have shown that some 50% of all TCS and TCC entering waste treatment facilities are adsorbed onto biosolids that are then released for agricultural use. (As this is written, perhaps half of all biosolids produced is the United States are used in agriculture.) In fact, the Targeted National Sewage Sludge Survey done by the U.S. EPA found that TCC and TCS were the top two pharmaceuticals and personal care products present in biosolids. And, of course, materials applied

in biosolids not only affect the microbial systems of the local fields where they are applied, but also wash off into surface waters and contribute to concentrations there as well.

Whether at some point higher concentrations of TCS and TCC in the environment create discontinuous or previously unanticipated effects, as in the case of CFCs, is as of this writing not known, but it certainly cannot be ruled out given the nonlinear behavior of these systems. For example, it is entirely plausible that ambient concentrations of these materials will enhance bacterial resistance not just in specific environments (the home or local doctor's office, for instance) but more generally across environmental bacterial populations as well. It is also possible that at some concentration such materials could trigger significant changes in the microflora and microfauna of critical ecosystems such as salt marshes, resulting in major shifts in populations.

So what are some of the issues for sustainable engineering raised by this case study? To start, this obviously reemphasizes the issues of scale, but with some subtle differences. Most obvious, of course, is the magnitude of the use and release of TCS and TCC themselves. Beyond that, such biologically active materials tend to react with certain metabolic pathways in the target organisms, as do many others. Some add to the toxic effect (act "synergistically"), while others might reduce that effect—an "antagonistic" result. Especially given the bioaccumulation potentials of TCS, TCC, their metabolites, and both synergistic and antagonistic loadings of other materials, it is apparent that, at least as a society, we can no longer address this on the level of individual chemicals or antibacterial products. As sustainable engineers, we now are faced with trying to understand the complexities of an emerging new earth system, one that includes land and water systems that have been essentially "redesigned" by the addition of antimicrobial and other anthropogenic chemicals.

Understanding this system is complicated by the fact that it is unlikely that any single entity will have all the relevant data. Disregard synergistic and antagonistic compounds (many not yet identified), and only consider all the products that may contain TCC and TCS. Some will be designed by chemical engineers working for cosmetic companies, some will involve soap companies, some will involve toy companies, and carpet companies, and toothpaste manufacturers. None of these firms will have much of an idea of what their entire sector is doing, much less the scale at which these chemicals are being designed into products in other sectors—much less about scale and production processes, and environmental chemistry and releases, of synergistic and antagonistic chemicals. Conceptually, of course, this sort of issue can represent a "market failure" in the sense that the information necessary to quantify costs and benefits, and thus lead to prices that include them all, doesn't exist. There

is thus a need for government regulation, but as we have already seen this sort of systemic regulation falls outside the institutional regulatory structure currently in place. Changing such institutions is very hard. Even if we could, it is not clear that this is the sort of simple system for which regulation is appropriate, or the type of complex system where other, more subtle, governance and incentive mechanisms might be appropriate.

Moreover, it is not clear that we have the ability yet as a society to balance the various values raised by such issues. For example, TCC and TCS, and many human drugs, are valued because of their action in the human body, usually maintaining health, addressing a disease or infirmity, providing entertainment (e.g., cocaine, residues of which are found in many urban rivers at high concentrations), or enhancing performance. Indeed, we design these chemical products precisely for such uses, usually tied to the particular products and services that a particular firm in a particular sector offers. Such regulation as there is to cover this lifecycle phase of a product tends to reflect these social and medical values: Does it work in the human body for the intended purpose? Is it more effective than alternatives, or no alternative at all? Does it cause more harm than good, especially considering nontarget organs? In other words, we evaluate pharmaceutical chemicals based on what they do in the body—a task that in itself is extremely difficult because of the complexity of the systems involved.

But come to the environmental impacts, and this calculus, complicated enough if one is just considering pharmaceutical issues, gets far more complex: How should one balance potential environmental impacts against real impacts of limiting a material that (presumably) has human benefits associated with it? How much poor health is a salt marsh worth? Such questions are not only heavily dependent on complex data issues (does the material really work?), but also point out a deeper difficulty. The answer to these sorts of questions is heavily influenced by the institution or group that gets to answer them. Deep greens will be more inclined to elevate environmental values over human health issues (for example, demanding a complete ban on dichlorodiphenyltrichloroethane (DDT) even when it is the only effective pesticide against malarial mosquitoes in some parts of Africa); medical professionals will be more inclined to value human health. Regulatory agencies will tend to reflect the worldviews of their constituents. Figure 9.4 illustrates some of these issues at different scales; especially at the final stage, there is a lack of data, institutional competence, and even robust theoretical frameworks.

Nonetheless, for the sustainable engineer it is clear that designing materials in today's globalized economy is often not just a matter of small volumes limited to a particular product, but is a matter of (unintentionally) designing broad earth systems as well. TCS and TCC are part of antimicrobial products, but they are also vehicles by which surface waters are being designed, and even less directly, by which ecosystems and aquatic communities are being

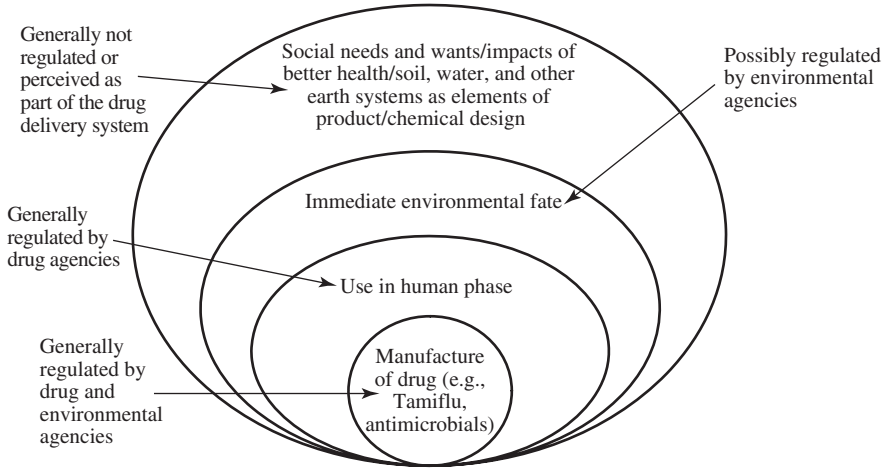


FIGURE 9.4 A Sustainable Engineering Perspective on Drug Design. This figure provides a model of sustainable engineering of pharmaceuticals and other personal care materials. The interaction of fragmented regulatory regimes with complex systems operating at many different scales in many different legal, disciplinary, and ethical domains, is one reason that engineering and policy practices have failed to keep pace with obvious changes in physical systems.

designed. We may not know how they do this, nor how to integrate such considerations into the complex economic, social, environmental, and institutional systems of drug and material design, but we can no longer ignore our responsibility for the fact that that's what is being done. In this case, it is the awareness of responsibility that is the beginning of sustainable engineering; the full challenge that lies before us is still to be taken up, not treated as already resolved. But simply a more energy-efficient way of producing TCS or TCC—a green chemistry approach—while useful, is obviously only a small, incremental, and potentially misleading understanding of true systemic impact.

9.4 POLICY IMPLICATIONS

This approach to green chemistry suggests a number of policy implications, which, for clarity, will be summarized at the bench scale, the material flux scale, and the earth systems scale.

As we have seen, the bench scale is the level at which much current work on green chemistry is being conducted. This is the level at which particular chemicals in individual applications are being evaluated and potentially preferable chemical processing methodologies identified. Much of this work is obviously quite useful, and the lesson from this scale is not that it shouldn't continue, but that the single-minded focus on this scale, within industry, the relevant professions, and education, combined with a reinforcement of this

bias towards the bench scale by green chemistry, is undesirable and unbalanced. As both the CFC and arsenic examples indicate, there clearly needs to be greater focus on the systems implications (especially regarding scale) that can be involved if an innovation proves commercially attractive. This is not just a simple matter of augmenting current focus. Rather, it implies a knowledge base that we currently don't have about the behavior of relevant systems at large scale, and indeed, a level of knowledge that we may never have a priori.

At this level, then, green chemistry is perhaps best thought of as good practice informed by heuristics with a presumption of environmental preferability. One would focus on such things as energy-efficient processes, minimizing use of toxics in processes and materials-use phases, and substituting less toxic materials where possible and all else is indeed equal. In many cases, green chemistry at this scale can have substantial cumulative benefits; if an application or material turned out to be popular, considerations of scale would obviously be appropriate. The idea here is that full systems analysis of the role of many materials is unrealistic, so that the theoretical best should not be allowed to become the enemy of the good that incremental implementation of green chemistry can achieve.

The next level can be thought of as the industrial level, where one uses tools such as material flow analysis to model and understand the flow of materials in commerce, as opposed to the instrumental level, which is more focused on bench-scale research and product or process development. Here, the issues are different, and the general knowledge state lower, so assumptions about what is preferable must be made more carefully. Private firms are important actors at this level, but they exist in a context of governmental regulation and, institutionally, are more frequently interacting with NGOs and other organizations with changing perspectives. In some cases, for example, the public may see firms that introduce new technology as having more of a social responsibility for the implications of the overall technology than the firms are used to assuming in a market-based economy (see Box 2.1 discussing the Monsanto experience with GMO technology). In other cases, unanticipated effects of materials over their lifecycle raise potential concerns, as in the pharmaceutical example. As a result, a sustainable engineering approach suggests that firms that think of themselves as chemical manufacturers might find themselves becoming material managers, responsible both to shareholders and to a broader group of stakeholders interested in systems that the firm's products affect.

At the highest level, of course, lie earth systems, which, it must be remembered, include not just the "natural" systems most people intuitively identify, but also "human" ones such as economic, technological, cultural, and social systems, and other built and designed systems, such as waste treatment systems. Here, the concept of green chemistry is least applicable, and, indeed,

can become dysfunctional, unless it is augmented by sustainable engineering approaches. For one thing, green chemistry provides little if any guidance in balancing environmental impacts, its sole focus, with social goods such as drugs that control dangerous diseases in people, or support high quality of life, or pesticides that save lives in Africa. Moreover, because of its focus on bench scale, green chemistry does not provide either the theoretical or methodological basis for the understanding, design methods and tools, and metrics appropriate to the overall system. This is an increasingly important gap because such a sustainable engineering framework will be increasingly important to navigate the challenges posed by nanotechnology and the material science and chemical engineering implications of the emerging technologies discussed in Chapter 8.

Obviously, then, green chemistry is a useful area of research, but far more limited than usually understood; this field has yet to either grapple with complex questions of social versus environmental costs and benefits, nor the scale and complexity challenge of earth systems, including modern globalized economic and technological frameworks. As the CFC example makes clear, chemical and material science innovations cannot simply be evaluated in snapshots: Unanticipated effects and perturbations may take time to develop and may be threshold phenomena, not occurring until a certain level of use is achieved. Thus, one area of research will need to be look at how the scientific and technical community can dialog with the relevant systems (or indeed, identify them) on an ongoing basis, and how system performance and state can be monitored. These are not just conceptual and research challenges; they are also profound institutional challenges. Developing mechanisms to monitor, dialog with, identify, and respond to the shifts in systems dynamics, and the ethical implications of human activity in shaping such dynamics, becomes a major challenge to sustainable engineering in this domain.

Exercises

9.1 Your government is considering a complete ban on the manufacture and sale of Bisphenol A within 5 years, based on the finding that it is an endocrine disruptor. This will affect the way your company manufactures plastic water bottles, baby bottles, and food containers, which form the mainstay of your production line. As the chemical engineer in charge of R&D as well as production, what steps will you need to take?

9.2 You are a sustainable engineer working for a pharmaceutical company. You have just developed a drug that is inexpensive to make and significantly reduces mortality from malaria. It appears, however, to be excreted in its original form, and is both stable and bioaccumulative. What concerns might this raise, especially given the areas of the world where you might expect it to be used?

9.3 Many environmentalists insist on a complete ban of the pesticide DDT, as it has well-documented negative environmental effects, especially on avian predators that are relatively high in the food chain. Many African countries in the malaria belt, on the other hand, believe that DDT-impregnated bed nets are the only immediate, inexpensive, and effective protection against malaria, especially in children. Who is right, and why?

9.4 Assume your job is to be the sustainable engineer in the public health system for the malaria belt. How would you construct a process that would enable an outcome that reflected all the relevant values, most especially those of public health for the poor, and environmental values on the part of global NGOs?

9.5 How would you redefine "green chemistry" so that it would be "sustainable chemistry"?

Annotated Bibliography

1. There are a number of books about green chemistry; perhaps the first and most influential is P. Anastas and J. Warner, *Green Chemistry: Theory and Practice* (Oxford: Oxford University Press), which first appeared in 1998. The literature on pollution prevention is related, and may be more practical for many; one of the best initial books in this area is the 1994 book from D. T. Allen and K. S. Rosselot, *Pollution Prevention for Chemical Processes* (Darby, PA: Diane Publishers).
2. Similarly, there are many books about the history of the chemical industry and related technologies; a useful one to start with is the 1991 text by F. Aftalion, *A History of the International Chemical Industry,* O. T. Benfey, trans. (Philadelphia: University of Pennsylvania Press). In addition, there is substantial information on particular materials available from the Geological Surveys and related professional organizations maintained by many governments, such as the USGS (which publishes the very useful *Minerals Yearbook*), although gathering and integrating the data necessary for such reports (and assuring that confidentiality of production of specific mines and firms can be maintained) means that the comprehensive data sets underlying these analyses are sometimes dated.

CHAPTER

10 Sustainable Engineering: Information and Communication Technology

10.1 INTRODUCTION

ICT is an important area for sustainable engineering for two reasons. The obvious one is that it is becoming a core technology not just in its own sector, but underlying every other sector as well. The automobile of today is so much more efficient than the 1960s muscle car primarily because of electronics that enable the substitution of information management for resource consumption (e.g., smart engine management to optimize gas mileage under whatever the sensor systems indicate is ambient conditions). The construction engineer and the civil engineer pack ICT into their houses and buildings, into bridges and road infrastructure, into water and energy management systems. The chemical engineer builds ever more productive and efficient factories because ICT enables systemic monitoring and control of conditions that would otherwise be impossible. Biotech engineers design and build new wetware—up to and including elements of the human—by using massive databases, machine searches, algorithms, and models drawn from electrical engineering and network theory. Crops are still biological, but increasingly developed countries deploy "smart agriculture" where sensors determine nutrient and moisture status of fields in real time. Meanwhile in our houses we design linked networks of appliances that can be operated remotely as needed. These appliances are hooked to a "smart grid" that manages them not just for our good, but to balance load across electrical infrastructures where both demand and supply changes can only be managed by creating highly complex information structures, which in turn raises serious and difficult security challenges as cyberspace becomes an active arena of military combat.

Digital natives roam real and virtual streets coupled to information appliances that tie them into networks like Facebook and Twitter, and open up engineering opportunities to build software castles in Second Life. Facebook in 2010 went over 500 million members, which, if it were a nation, would make it the third largest on Earth in population, behind only China and India. That virtual reality and social networking sites, and nations, are similar is not as far-fetched as it might seem, given changes in both types of institution. *The Economist,* a respected journal, in a 2010 article titled "Social Networks and Statehood: The Future is Another Country," noted that "despite its giant population, Facebook is not quite a sovereign state—but it is beginning to look and act like one."[1] And the most successful retailer of our day, Walmart, competes based more on the effectiveness and competence of its ICT infrastructure than on its products, which are inexpensive in part because Walmart is so efficient in its operations. That efficiency in part means Walmart can operate on virtually no unnecessary inventory, which reduces the significant if hidden energy and material cost associated with inventory in retail pipelines that never gets sold and must be discarded or recycled. Whether this constitutes sustainable logistic system engineering no doubt depends on how one thinks about current consumption patterns, but its efficiency is not in doubt.

More subtly, however, ICT has been an underlying infrastructure and integral part of the continuing evolution of the human species for millennia. Verbal capabilities date to the dawn of being human and derive from evolved intellectual, biological, and perhaps cultural competencies. But the invention of writing, of books, of the printing press, and technologies such as the telegraph, radio, and television, were all critical enablers of cultural development and diffusion. On the pictorial side, art began the process of both shaping and retaining the past, and led to its own technological evolution (such as new and more vibrant paints); the evolution of photography, then "moving pictures" and "newsreels," meant that "real life" could be captured directly. Digitalization technology and cyberspace now mean, of course, that no image can be trusted and that memory itself is in play in our technologies as well as in our bodies, because drugs are being developed to manipulate memory directly in the brain as well. ICT and cogsci are combining to make the past as explicitly constructed as it frequently was implicitly. (Any good photographer, even before Photoshop, knew how to take an image that projected the appropriate emotional message.) The evolution of modern mass media, combined with the growing power of the Internet, however, has created a new and powerful technological and cultural dynamic, and enabled social and cultural tides that are, for the most part, not even yet perceived. Like the global transportation network that evolved as European civilization spread around the globe, the

[1] *The Economist,* July 24, 2010, pp. 59–60.

Internet will have huge impacts on the way humans manage the physical and biological world—and the way that world, increasingly endogenous to human society, evolves. Indeed, the Internet has become an important mechanism by which humans perceive their world, including phenomena that would be impossible to detect otherwise. No one, for example, has ever directly seen or felt climate change; however, massive models tell us it's there, and the Internet not only facilitates the community that supports those models, but spreads that message worldwide. The fundamental dynamic cannot be ignored. Our ICT technologies are increasingly the mechanism by which we see—indeed, construct—our Earth. As a corollary, understanding ICT thus becomes an important part of sustainable engineering.

10.2 FRAMING ICT

An important part of sustainable engineering is learning to deal with wicked systems that include meaningful human and social dynamics at many scales. These systems often present themselves to the sustainable engineer without any useful differentiation. Accordingly, an important skill is being able to map systems for purposes of analysis. It is not so necessary that what you produce be "right" as that it give you a framework within which you can work.

ICT presents such a complex mixture. Even at the most prosaic level the industry itself is in a period of rapid economic and technological change, with sector boundaries and core technologies undergoing fundamental transitions. Digital cameras replace film cameras, and in turn are in the process of being replaced by cameras as a mobile phone function (as that device incorporates different functionalities, is it still a mobile phone?). Moreover, there are serious technological tipping points that have yet to settle out. Will broadband access be provided primarily through hardwired or wireless facilities, and does that answer depend on the jurisdiction involved? Video and audio entertainment, which used to be provided on storage media (tape, CDs) in specialized locations (e.g., movie theaters), or in broadcast mode (e.g., over-the-air or coaxial cable based networks) is now increasingly available in all forms over the Internet on demand; moreover, because much of this content is being provided without necessarily following the niceties of copyright law, the economic implications for existing industry structures are profoundly disruptive. National and regional newspapers, the prototypical mass media outlet, are in serious trouble and are being displaced by blogs and other sources, with interesting implications for content and the way societies jointly construct their versions of truth and reality. (Local newspapers may be less threatened, although this is a sector in turmoil, so little can be predicted confidently.) In addition, many of the sustainability and environmental implications of ICT arise not from physical platforms, facilities, and artifacts, but from the services

provided. Social networking, telework, and e-commerce, for example, can be anticipated to have far greater social, commercial, and environmental effects than those created by the underlying physical platforms. Such "direct" services, moreover, are dwarfed by the implications of building intelligence into the operations of virtually all sectors (as in augmented cognition, or "augcog," discussed in the next chapter). Finally, there is the entertainment and content industry, with profound but subtle effects on attitudes and worldviews.

As with every complicated system, there are difficult boundary questions. For example, there are at least two critical information systems that must be recognized as such, yet are usually considered beyond the scope of a discussion of ICT proper. One is the market, which is one of the most powerful and remarkably effective information systems on the planet (regardless of where one stands on it ideologically). The second is the genetic, proteomic, and other higher level information structures that are the province of biotechnology and genomic engineering. Interestingly enough, much of the biological information system is transitioning to huge databases, and from thence to intellectual property, and from there into the market.

Obviously, one chapter cannot cover the ICT landscape in detail; indeed, there are many issues, such as the psychological, cultural, and institutional implications of social networking, which are not yet understood well enough to talk about even if we had space. But, following the sustainable engineering principle of addressing a complex and inchoate system by first constructing a working conceptual framework, let us begin by building such a qualitative schema.

Because we are interested primarily in the engineering aspects of ICT, we will take an approach that focuses more on technology rather than on information content. Focusing on content would give us a very different perspective on ICT; that is not a weakness of this particular approach, but rather an indication that complex systems are virtually never completely understandable through just one approach. In that spirit, Box 10.1 presents some elements of a firm-based perspective, which is also different.

BOX 10.1 A Telecom Firm Perspective on Sustainable Engineering

As mentioned in the main text, any complex adaptive system such as ICT can be looked at from many perspectives: technological and engineering; cultural; content; individual; institutional; and many others. Each will highlight certain aspects of the system, and downplay or ignore others. A technological approach, for example, would not emphasize the cultural implications of social networking, or the notable shift towards violence and visual impact in modern cinema. Whether a particular perspective is "right," then, depends not on how complete it is, for all perspectives on a CAS are partial. Rather, it depends on the reason for which a particular perspective

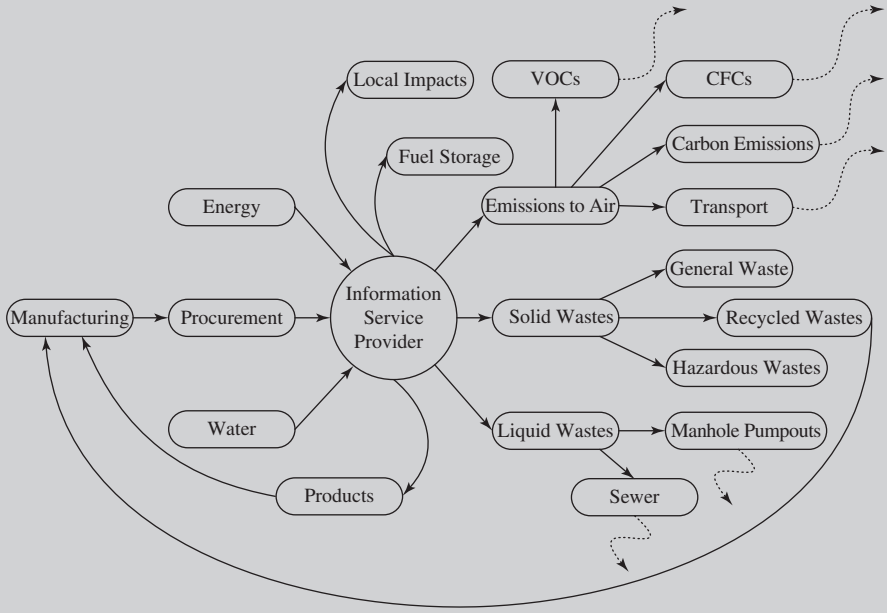


FIGURE B10.1.1 Information Service Provider Environmental Lifecycle. This figure illustrates some of the environmental impacts of a telecommunication firm's operations. When quantified, even high-level maps such as this can provide information enabling the firm to improve the environmental efficiency of its operations, often at a cost savings as well. "VOCs" are volatile organic compounds—short-chain hydrocarbons such as chlorinated solvents and alcohols.

(Source: "BT and the Environment: A Report on BT's Environmental Performance 1996/1997." London: © British Telecommunications plc, 1997. Reprinted with permission.)

is chosen. If one is interested in sustainable engineering, therefore, one would probably find the technological perspective most useful, at least as a start. If one is an executive at an ICT firm, one might, however, want to take a different perspective. Consider, for example, Figure B10.1.1, based on a telecom firm annual report.

Suppose, for example, that one were the Vice President of Environment, Health, and Safety for a large telecommunications infrastructure firm, and thus interested in the environmental impacts of the operations of that firm from a systemic perspective. Such a perspective suggests certain boundaries and constraints. Assuming that the firm is a pure service provider (i.e., it does not manufacture its own gear), it is unlikely that it would extend its analysis very far down the supply chain absent some sort of pressure from customers or important stakeholders. A large firm will have tens of thousands of suppliers, many of which will have their own, similarly sized, network of suppliers, many of which will change unpredictably and variably as components and materials are purchased on the open market. Accordingly, the returns from a deep supply network search are expensive and

of limited value. It is also doubtful that, without some reason, one would go too far beyond looking at basic services. In both cases, it is also important to note that the firm will have dramatically less expertise the further it moves from its core operations. Thus, when the author was with a telecom firm, the practice with critical suppliers was to ask them to show that they had considered social and environmental factors in their design, but not to try to second guess their conclusions, on the grounds that a service firm didn't know enough about the intricacies of manufacturing to do so.

A private firm perspective also assumes that government policies are by and large a given. Some may be onerous, some may not make sense, and in most cases any large firm will be in constant dialog with various government jurisdictions over many issues, but policy changes will not be a big factor in the perspective simply because of the lack of control. Conversely, more attention will be paid to internal standards and practices than might be justified by a different perspective, because those the firm can control. Thus, for example, it is unlikely that "manhole pump-outs" (draining potentially contaminated water from manholes that one needs to work in) will be of serious concern on a national environmental basis, but to the firm, they are both an important potential liability to manage and within the complete control of the firm.

Finally, note the environmental emphasis, and recall the point frequently made throughout this book: Green engineering is an important part of, but is not coextensive with, sustainable engineering. Firms are more likely to focus on environmental issues because they are more regulated, more evident to consumers, and easier to understand and communicate. However, as this chapter makes clear, these effects are probably the least important, in a systemic sense, of the many sustainable engineering implications of an ICT firm.

Taking the technological approach, therefore, we can parse ICT into six components, each with its dominant level of technology and particular considerations from a sustainable engineering perspective (Figure 10.1):

1. The materials required to build the artifacts that underlie ICT, from mobile phones and flat panel displays to routers and switches
2. The manufacture of such artifacts
3. The lifecycle of such artifacts
4. Construction of the networks that, using a wide variety of different technologies, including both hardware and software, enable delivery of ICT functionality
5. Services offered over ICT systems
6. The evolution of the infosphere that increasingly provides humans with their entertainment, perception, and understanding of their society and their world, and even identity as humans

Technology Category	Level of Technology	Analytical Approaches	Sustainable Engineering Considerations
Artifact material sourcing	Level I (low cultural content except for analysis of source risk)	Market signals; scarcity and source stability analysis	Does the material work; is it freely available at full production demand levels?
Artifact manufacture	Level I (low cultural content)	DFE, manufacturing process design issues	Energy efficiency and toxics reduction in manufacturing processes; worker safety
Artifact over lifecycle	Level I and II (use patterns depend on services and local culture)	DFE, lifecycle analyses; marketing and business plans	Design for foreseeable conditions of use; energy efficiency in use; end of life management; user safety
Construction and maintenance of networks	Level I and II (network design Level I, but design objectives and constraints depend on local conditions, services)	Systems engineering; cross-infrastructure integration	Efficiency, stability, and security of network and critical nodes; proper operation; managing technological change
Services	Level II and III (large cultural component)	Combination of marketing and engineering	Both short and long term effects; difficult to predict success a priori; services and technology often tightly coupled
Evolution of infosphere	Level III (unpredictable, earth system)	Earth systems engineering and management; integration of engineering and social science analyses	Real time technology and policy assessment required; large unpredictable shifts possible

FIGURE 10.1 The Information and Communication Technology System. This schematic presents some ideas about sustainable engineering considerations at various levels of the ICT system. This framework reflects a sector scale technological perspective; one based on the content industry, or Internet services, or the issues faced by a particular firm, would look quite different.

Note that many other divisions would work equally well; what we are after is a workable framework, not the "best" framework. Trying to find the latter is a good guarantee that we'll never get anywhere. Moreover, although this division enables an understanding of the gradation from Level I technology in ICT to Level III, many of the tools available cut across different technology categories with greater or lesser efficacy. For example, sustainable engineering questions at the manufacturing stage, such as using the least toxic but still effective solvents

possible, are pretty clearly Level I. A good choice carries with it control of worker safety and environmental issues.

Looking at the product over the lifecycle, however, becomes more complicated. For example, one can reduce the energy consumption of electronic cell switches by increasing the transmission energy of individual mobile phone units, but that would consume significant power from consumer units and might have health effects if the increased power had some sort of physiological effect on humans (a contested theory at this point). Moreover, the use phase of a product and its design is strongly determined by the culture within which it will be used, which in turn brings different design objectives and constraints into play, depending on the particular market in that culture for which an artifact is being made. This is Level II territory—understandable, but with complex intermingling of technology, culture, environment, and economic domains.

Especially in the electronics sector, tools have been developed at the Level I and Level II stages. Such a tool is Design for Environment, or DFE, an approach that has been mentioned previously, developed over the past fifteen years or so primarily in the ICT sector. Exploring DFE in a little more detail is useful not just in itself, but also as an example of how some tools are particularly useful at certain levels of a system, in this case the product manufacture, design, and lifecycle stages. It is also helpful to an understanding of sustainable engineering to see how such tools evolve over time (see Box 10.2). In this case, fifteen years of history lie behind a tool that, because it focuses on environmental issues, is only a partial component of a sustainable engineering approach—but there is a historical reason for this bias, and it is a bias that broadly colors sustainable engineering even today.

With services and the infosphere, we clearly enter Level III. What, for example, is the "goal" of social networking, and what will its eventual cultural effects be? What will happen when the market, the Internet, and biological information systems begin to merge? These sorts of questions are best addressed through development of relatively simple scenarios used to stimulate discussions and option planning (Box 10.3), but it is clear that, regardless of specifics, the directions in which ICT is driving humanity are beyond our ability to predict or understand a priori. At this stage, our toolboxes, and even our understanding of the relevant issues, are pretty much empty.

10.3 INFRASTRUCTURE OPERATIONS

Virtually every product implies an infrastructure. Home appliances require an electric infrastructure; cars require gasoline and roads; airplanes require airports and jet fuel. In the case of ICT, there are a number of infrastructures in addition to the obvious operational ones consisting of landline and

BOX 10.2 Design for Environment

DFE is now a widely used term in sustainable engineering and industrial ecology, first introduced in the electronics industry in the early 1990s. As a module of the Design for "X" process discussed in Chapter 6, which was widely used in the electronics industry at the time, it was seen as a mechanism for driving environmental considerations and constraints into the product design process. Although it reflected a number of then-new initiatives such as "green engineering" or lifecycle assessment, DFE thus began as a module for a specific design process and was only later broadened to become a generic set of practices for integrating environmental considerations into the design of complex products such as computers or telephone switches.

A short history of the process by which DFE was institutionalized is useful as a case study of how to implement change in industrial practices. The intellectual basis for DFE and its linkage with industrial ecology was first established in a dissertation completed in 1992.[2] Creating such an intellectual foundation was important, in part to shift environmental issues from being merely incidental to manufacturing to being strategic, but in itself was entirely inadequate as a force for real change (that is, for adoption of such practices by industry, in ways that made a difference to their operations and their product design). The institutionalization of DFE, accordingly, began with the formation of the Design for Environment Task Force at the American Electronics Association (AEA) in 1990 and the subsequent publication of a set of ten "AEA DFE White Papers," which introduced the broader electronics industry to the concepts of DFE. Other efforts drawing on this material extended the institutionalization process; for example, the U.S. EPA subsequently formed a DFE group at the Agency, with a broader remit than just ICT. Other industry groups, such as the Microelectronics and Computer Technology Corporation (MCC) and Sematech, issued further reports and road maps exploring in depth the material, energy, environmental, and management issues raised by electronics products and components (e.g., integrated circuit packaging, printed wiring boards, displays) over their lifecycles. The broad industry participation, and the fact that those involved were primarily involved in research and development, and manufacturing engineering, rather than in environmental and compliance activities, made these documents both technically rigorous and a significant departure from initial practice, which had too often seen environmental professionals trying, with a notable lack of success, to prescribe to design and manufacturing teams. An important source of continuing institutionalization involving technical, rather than environmental and safety, professionals was the annual symposia on electronics and the environment, focusing on DFE, held by the IEEE from 1993 to the present. Indeed, the Proceedings from these symposia are perhaps the single greatest resource for exploring the progress and elements of DFE as it has evolved for the electronics industry.

---

[2]Brad Allenby, "Design for Environment: Implementing Industrial Ecology," Ph.D. dissertation, Rutgers University, 1992.

From about 1995 on, however, an important jurisdictional shift gradually occurred. Although initially organizations centered in the United States, such as AEA and MCC, had led in research activities, the growing European interest in product takeback, especially as applied to electronics, led to a growing intellectual dominance by European research initiatives. Response by electronics manufacturers was global, but there was little question that the primary driver of technological evolution shifted from American waste reduction and remediation policies to European product takeback, material ban, and lifecycle product management initiatives.

For the sustainable engineer, this case offers several important lessons:

1. Try to minimize the apparent impact of any change. For example, had the individuals involved in crafting DFE tried to implement a completely new product design process or developed a tool incompatible with what was then current practices, it is highly doubtful that they would have been successful. Instead, the path chosen—to create another new module in an existing process—made the desired change easy, almost invisible, to design teams, and for the most part they were glad to cooperate.
2. Work through the appropriate organizations. DFE introduced green design principles through the design organizations, rather than through the environmental organization, and thus gave managers, manufacturing engineers, and product teams the comfort of knowing that design professionals had vetted the content and processes, and that they were therefore likely to be useful and valuable in product design. In contrast, attempts made by a company's environmental group to intervene directly with design suggestions generally failed. In retrospect, the reasons were obvious. The environmental professionals, although respected for their environmental knowledge, were not viewed as either understanding, or being able to usefully participate directly in, design activities (a perspective that, at least in many cases, was accurate).
3. Build cross-sector buy-in by using appropriate trade and business organizations. A major reason for the success of DFE was its early discussion and acceptance by such respected professional organizations as the IEEE, the AEA, and others. The more subtle message here is to work not just with your peers within the firm, but with those who agree with you in other firms. In today's complex commercial environment, firms and their professional staffs both compete and cooperate, and when developing useful basic tools such as DFE, the competitive dimension does not rule out a great deal of collaboration in establishing practices, methods, and a community of professional practitioners.
4. Early in the institutionalization process, create the necessary intellectual frameworks and publications around which practices and communities can be built. For instance, although they were relatively simple and are not unjustly forgotten today, the AEA White Papers mentioned above were an important step in institutionalizing DFE; analogously, creating the *Journal of*

Industrial Ecology was an important and deliberate step in institutionalizing the field of industrial ecology.

5. Note the important role that industry played in building DFE. This is in one sense only to be expected given that the issue area is product design, which is an industrial, not a governmental or NGO, competence. On the other hand, this was quite new at the time, so it indicates the need for the sustainable engineer to be flexible and nondoctrinaire about where the best opportunities to do good engineering may lie. For example, the primarily public effort involved with engineering the Everglades to remain viable for the long-term is an ideal opportunity for sustainable engineering, even though the involvement of the private sector is secondary.
6. Note the historical progression: DFE begins with the shift of environmental issues from being merely ancillary to being strategic to the firm (product and process design). It then begins to take a more systemic perspective on environmental issues, adopting a product lifecycle perspective. Finally, DFE is now the inspirational and methodological basis for the early stages of developing Design for Sustainability (DFS), which adds the social and cultural domains to DFE approaches. (The economic aspect by and large is taken into account because DFE occurs in an industrial context, where economics are part of virtually every routine and decision process.) At this point, however, sustainable engineering methods are not yet adequately differentiated from green engineering or DFE approaches, indicating that much more thinking and work are required to develop DFS tools, frameworks, and methodologies. The lesson? Draw from as many sources as you can—Design for Reliability, for example, to extend product life; or Design for Manufacturability to reduce waste from the manufacturing process—but be careful to always remember that sustainable engineering will require more than just borrowing existing technologies.
7. Be prepared for globalization to throw curve balls. In this case, much initial DFE work was done in the United States, especially by American electronic manufacturers, but the center of manufacturing shifted over the next decade to Asia, with the result that DFE was no longer nearly as important to American firms, who were in turn moving into primarily service provision rather than platform manufacturers. Moreover, the center of environmental regulation and legislation shifted over to Europe from the United States, with the result that European policy initiatives such as product takeback in various forms became leading edge. The net result is that the United States is far less important to the continuing evolution of DFE and related practices such as LCA and the field of industrial ecology than it was ten years ago. The same is not true, however, of sustainable engineering, in large part because the United States remains perhaps the single most important source of technological innovation on the planet—and sustainable engineering is deeply engaged by such innovation.

wireless networks. Additional infrastructure includes office buildings and vehicle fleets, to more specialized requirements such as extensive battery back-up systems for switches and routers to provide service continuity if the electric grid goes down, to less material infrastructures such as the financial and software systems that enable operation of the physical infrastructures. There are strong financial incentives to optimize infrastructure operations, such as reducing energy consumption resulting from network operation, but there are also significant constraints in doing so. To take one example, a service operator has little control over the energy consumption of the routers, servers, and switches it buys, although it can reduce energy consumption somewhat through innovative network design decisions. Similarly, many operators of server farms are exploring moving them to northern climates because a major energy expenditure for server farms is keeping them cool, which is, of course, less costly when the ambient climate is on average cooler.

An important step in any such infrastructure efficiency evaluation is to have some basis for prioritizing activities. This has long been industry practice. To take a historical example, British Telecom's (BT's) 1996–1997 annual environment, health, and safety report, among other things, tried to take an initial look at the relative environmental importance of certain infrastructure activities compared to others (see Box 10.1). Their data indicated, for instance, that network and infrastructure operations, as opposed to office building and fleet operation, constituted the single biggest source of emissions, contributing around 90 percent of carbon dioxide and 98 percent of sulfur dioxide emissions. Similarly, back-up power in network operations proved more important than fleet battery requirements; the preponderance of batteries recycled were lead acid network exchange batteries (99 percent in 1996, 87 percent in 1995, and 93 percent in 1994), as opposed to vehicular lead acid or NiCad batteries. And, not surprising given this period was seeing major changes in network technology, the major metal containing residual stream from ongoing operations was from network switching exchange construction and replacement activities (constituting 65% of the total), as opposed to battery or cable recycling, or other such activities. What all of these data pointed to was that, from an overall systems priority, the most important focus in terms of environmental efficiency was the network infrastructure rather than other operational infrastructures such as vehicle fleets. An additional benefit of this prioritization is that network technology and design is more under the control of the firm doing the analysis than, say, the fleet technology. A telecom firm can streamline its own operations a lot more easily than it can demand redesign of fleet vehicles (it obviously has far less technical competence in doing the latter, as well).

These sorts of systematic infrastructure data are obviously quite useful to the sustainable engineer: They help define the system quantitatively, thus aligning with engineering culture and practice; they identify major as opposed to minor material flows (and thus potential impacts); and they point towards priorities for management of environmental impacts. Remember, however, that these particular data are environmental, and thus provide only part of the information required for sustainable engineering, a particularly important gap with ICT sectors because they enable efficiencies not just in themselves, but across the economy. Moreover, the risk is that, because environmental data are relatively easy to obtain and quantify, they will be used as proxy for sustainability data. The latter, if done comprehensively, would tend to be far more balanced in identifying costs and benefits. This is because virtually any industrial operation negatively affects environmental systems, and so is highly likely to indicate costs rather than benefits for any particular operation. This does not make environmental data "wrong," only partial, and with a necessarily consistently negative bias.

10.4 SERVICES

It is at the level of services that the first faint traces of Level III complexity begin to intrude. The alert reader will have noticed that, until now, the primary methods suggested as models for sustainable engineering are environmental and have a relatively long history behind them. It is at the point where one must deal with services that such historical precedent has tended to wither: Traditional environmental studies and policy, as well as related fields such as industrial ecology, tend to exclude services from consideration. The reasons are complex, but include the traditional environmental focus on solid waste, air and water emissions, and material and energy consumption, all of which are much more obviously engaged by products and manufacturing rather than services. Moreover, manufacturing is done by firms, while many services are consumed directly by individuals. It is thus politically much easier to constrain the former rather than the latter. (So, gas efficiency standards are applied to automobile fleets, and thus to car manufacturers, with far less political cost than gasoline taxes are imposed on the public.) Nonetheless, it is clear that, especially with a sector such as ICT, any analysis that stops with the artifact is profoundly unsystematic and certainly not adequate for any type of sustainability analysis. When someone buys a DVD of a movie, focusing on the plastic in the DVD rather than the content of the movie completely misses the principal impact of the transaction. When a digital native goes on a social networking site, to worry about the energy consumed in the process as opposed to the cultural and psychological implications of the service is to profoundly misunderstand the sustainability of the system. It is not that the

material and energetic implications of the DVD or electricity used in computing aren't meaningful; it is simply that, for sustainability purposes, they are a very small tail on a very large dog. In particular, it is important to remember that ICT services provide a great deal of quality of life to many people in many different ways, and thus those benefits must be included in any sort of sustainability analysis of ICT.

That said, it is certainly true that, although our understanding of the psychological, social, cultural, and environmental dimensions of ICT service areas such as digital nomad working patterns (the separating of work from location, including virtual offices, telework, or working from the local coffee shop), or social networking, or Internet entertainment delivery over portable communication devices, or virtual reality spaces and games, is improving, we are far from comprehensive or even adequate understanding. In part, this is because of basic psychology, particularly the human inability to understand the future except in terms of the past, a limitation that becomes especially potent during times of rapid technological evolution. It also reflects the fact that, with any suitably powerful technology, it is impossible to predict how it will be adopted, and coevolve with, social and cultural institutions going forward. This is particularly important in the information industry, where the rate of technological change is unprecedented. This has been true for some time: Maximum Internet trunk speed was doubling every twenty-two months, as was maximum router speed—at least until 1997, when the doubling period shrank to six months. Internet traffic growth doubled every twenty-one months from 1969 to 1982, but the doubling rate shrank to nine months between 1983 and 1997, and to six months after that (in part reflecting a shift in traffic patterns from words to music, photo, and video files). Financial trends have been similar for a while: the U.S. Internet economy generated $301 billion in revenues in 1998; but more important, from a "future shock" perspective, it was growing at the time at a compound rate of about 175 percent per year at a time when the U.S. economy as a whole was growing at about 3 percent a year. Such dramatic growth rates are not only continuing, but are difficult for people to process because they lie outside our normal experience.

An obvious way this evolution affects sustainability lies in material use. Consider that software distribution, whether operational software or entertainment, used to be via strictly material means: DVD, CD, tapes, and (for those of you old enough to remember them) floppy disks. The past few years, however, have seen a substantial shift in technology platforms to a dramatically more interconnected and useful set of physical networks (cable, wireless, mobile phone, satellite), all of which have an increasingly powerful and functional Internet platformed on them. The result is a radical, if little noted, dematerialization of information (and because transmitting information is far less expensive than even the small amount of material in CDs or DVDs, a

significant drop in cost). This does not just involve the storage artifact itself, but also the packaging (often quite extensive), the physical transportation and mailing of such artifacts, and the entire built infrastructure necessary to support the business model of material-based media delivery. The result is in some ways classic sustainable engineering: Quality of life, in the form of on-demand entertainment and information of any kind, available at any location, is enhanced with far less environmental impact. On the other hand, this transition obviously involved economic costs (and still does) to existing businesses, shareowners, and employees. Such transition costs are not unique to socially and environmentally preferable technological evolution, but are characteristic of constantly changing capitalist economies in general. In addition, the social benefits, although no doubt real, are not unmitigated. There are always questions of equitable distribution of costs and benefits when such sustainable engineering technologies evolve: What happens to the workers in DVD or CD manufacturing facilities or video stores as they are made redundant? How rapidly do industrial organizations adapt to the change—or like the record industry, do they pretend technological change is not affecting them, and thus by choice and inertia become yet another unnecessary casualty of change? Moreover, one can envision so-called rebound effects: Is it not possible, for example, that on-demand delivery of Internet-based entertainment and information to automobiles will simply make traffic jams more pleasant, thereby reducing incentives not to get caught in one?

In regard to this example, note that the changes involved tend to be systemic and unplanned, rather than under the control of any individual or institution. (Indeed, to the extent such changes can be controlled by existing institutions, it will often be to the end of stifling them, rather than encouraging new technologies and practices.) This implies that an important sustainable engineering skill is learning to discern such changes in underlying systems in order to take advantage of them, rather than actually being able to control or design them in any detail. A possible exception is the situation where one institution can establish standards regarding a new or evolving technology, and thus favor a particular design; companies such as Microsoft have grown rich from doing so.

As this (now historic) example suggests, an immediate result of accelerating ICT evolution has been the increasing inability of governments, institutions, academic research, and even individuals to keep up with what the service dimension of ICT can do and what it means. Augmented reality, where mobile phones provide real-time Internet overlays on physical landscapes, is a commercial reality as well. Many firms predict that "Turing Virtual Reality"—where virtual landscapes will be so effectively rendered that an individual, placed in front of two flat panels, will be unable to tell which is a window and which is virtual (except perhaps for the elves)—is a decade or less away. A core

competency in sustainable engineering is designing appropriate ICT capacity into the artifacts or systems being worked on, from weapons platform to urban infrastructure to buildings to cars. These capacities not only change how systems behave in unpredictable ways, but begin to distribute learning capacities across landscapes in ways that promise profound change. (What does a battlefield look like when operations are being overseen and coordinated through networked "smart dust" nanoparticle sensors and computers, augmented by grid computing cyborg insectoids?) The rapid pace of technological evolution, and the significant psychological and cultural dimensions involved in information production and consumption, makes generalizations at this point superficial.

Under such circumstances, the tools of choice for sustainable engineering are not predictive, but rather methods such as scenario development and analysis that enable one to play with possible futures. Developing and using scenarios is an important sustainable engineering skill, because it is one way to deal with conditions of high unpredictability, uncertainty, and complexity. It enables better responses, but does not pretend to predict the future. (Box 10.3 provides more information on scenarios in planning and design.)

10.5 THE VIRTUALIZATION OF WORK, DIGITAL NOMADS, AND THE TRIPLE BOTTOM LINE

To explore some of the implications of services, let us use one example, the growing practice, enabled by ubiquitous wireless technology, of being able to perform knowledge work anywhere. This has created a class of individuals, sometimes called "digital nomads," who carry their ICT presence with them and participate in what might be generically termed "the virtual workplace." Occasionally this may indeed involve meeting or working in a virtual reality such as Second Life; more often, it is a simple matter of working from the home office or a nearby coffee shop rather than driving for an hour to a central office building. Figuring out how to enable digital nomads so that they are more productive (economic benefits), commute less (environmental benefits), and happier working (social benefits) is nontrivial sustainable engineering. Moreover, it is only the beginning. You start with what appears to be a relatively small systemic change—say, designing the ITC infrastructure for a firm considering more virtual work emphasis—and you end up redesigning the idea of the firm itself.

This brings up an important point. Change is often both incremental and discontinuous, and it is important to understand the difference. Thus, with virtual office initiatives, there are two very different situations: a virtual office situation where a few days a week are worked nonlocally but most work is done from a central location, and another where one works from anywhere

BOX 10.3 Scenarios and Scenario Planning

One of the most difficult challenges with complex adaptive systems and unpredictable evolution of emerging technologies is very basic: Given that, by definition, I don't know much about these systems, how can I design or manage them? The answer is: Generally, you can't directly design or manage them, but by practicing on possible futures, you can develop the capability to react to them as they evolve in ways that are reasonable, rational, and ethical. In other words, you must practice the skill of adapting to the future in real time.

The best way to do this is the strategic planning method called **scenario planning** or **scenario analysis**. Used widely by military organizations because of the profound uncertainty and rapidly changing environments within which they operate, scenario planning has been widely adopted by corporations, beginning with Royal Dutch Shell in the early 1970s. Scenario planning grew out of simulation gaming, where known facts about the future, such as geographic, material, political, or military conditions, are combined with selected alternative hypothetical conditions to create a future that is at once both rich enough to challenge and explicitly understood to be hypothetical rather than predictive. The fundamental difference between a scenario exercise and a prediction is that scenarios are explicitly hypothetical; predictions pretend to be real visions of actual future states. Any statement in this book or any other about future technological states, for example, should be understood as a scenario—maybe a high-probability scenario, but a scenario—rather than a prediction about "what will happen." This is obviously a mistake that is common with many modelers and those who interpret their results. Models produce (hopefully high-probability) scenarios given certain assumptions; they do not produce predictions of actual future state.

Scenario planning is most useful where the number of variables and potential changes make simpler approaches less effective. (If you're worried about incremental changes or shifts in only one or two variables, there are other planning and gaming techniques that are preferable.) The essence of scenario planning is creating a family of scenarios that diverge by making different assumptions about highly uncertain and deeply influential driving forces—Royal Dutch Shell, for instance, used the price of a barrel of oil, among other variables. These are then discussed and evaluated in a group process, which tries to draw out the various implications that the scenarios raise, and suggest possible responses, all in an admittedly hypothetical space. Then, should events begin to play out along the lines of a scenario—say, the price of a barrel of oil increases dramatically—the institution or firm is in a much better position to respond than if it were caught completely by surprise. Scenario planning at its best enables managers, executives, officers, and policy makers to make mistakes and learn from them, and explore preferable options and routes, without having to do it in real time under real conditions where mistakes are often costly and even disastrous.

Although the specifics of scenario planning differ depending on the institution and the type of challenge faced, there are a few steps that most scenario planning exercises usually include:

1. Decide what the problem and thus the scope of your scenario should be. Ideally, a scenario is temporally and topically broad enough to be interesting, without being so broad that important parameters are essentially unknowable. In practice, this usually leads to scenarios looking at a ten-year time frame, plus or minus.
2. Identify (a) major stakeholders, and (b) the team that you want to be involved in the exercise. These two groups may or may not overlap at all.
3. Determine the key driving forces, looking particularly for variables that have high levels of influence on system behavior and performance, and are highly unpredictable. There are a number of techniques to help you do this, such as mapping driving forces on a compass graph with one axis being predictable/unpredictable and one being important/unimportant, then looking at the items in the "unpredictable/important" quadrant.
4. Develop a relatively small number of appropriate scenarios, which can be fleshed out by including other driving forces and variables, additional stakeholders, and the like. The planning group should be part of this process, which should be free to create additional scenarios if necessary to explore option spaces and to drop ones regarded as ineffective.
5. Evaluate the implications of each scenario for the institution and determine reasonable responses as the scenarios play out. This is often the most important part of a scenario, as "business as usual" seldom works when changes are discontinuous and severe, rather than just incremental, and the process of even having to consider them can substantially improve the adaptability and resilience of an institution.

Some elements of sustainable engineering involve the application of well-known heuristics—"rules of thumb"—such as designing for energy efficiency or reducing unnecessary use of toxics. The more complicated dimensions of sustainable engineering, however, frequently involve trying to "future proof" designs, infrastructure, products, and systems—that is, trying to make them resilient against unpredictable, and likely very difficult and fundamental, challenge. Traditional tools are often not helpful in this process, which requires developing technological, institutional, and policy options that enable responses to the unexpected. Scenario planning, which helps support development of systemic resiliency in complex adaptive systems with significant human components, is thus an important and useful method for the sustainable engineer.

(essentially, in cyberspace) (see Figure 10.2). People are often very comfortable with the first situation, but not with the second; indeed, the second raises far more fundamental organizational and technological challenges. Occasionally working at home does not require major changes in corporate ITC infrastructure, for example, because the relevant work can simply be taken home physically.

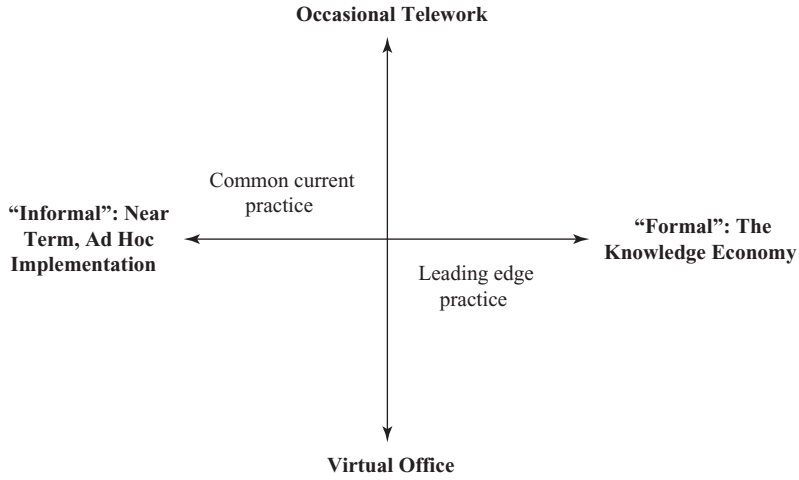


FIGURE 10.2 Virtual Office Compass. This figure maps the difference between incremental change to existing practices, and the formal embrace of the knowledge economy, with all its implications. Many firms are doing incremental virtual work, but very few are truly flat, networked, virtual organizational structures existing mainly in cyberspace. In part, this may be because for many firms, some of their operations may lend themselves to virtualization, while others—running a manufacturing line, for example—may be harder to virtualize entirely. In addition, elements of hierarchy are important for effective operation as a firm becomes larger and more complex.

But to support a completely virtual operation requires that not just the ICT system (including critical components such as network security) be up to the job, but that the information required for work be available in cyberspace. In other words, designing and managing a virtual office initiative is not just a technical engineering task, but also a deep institutional challenge. This caught many firms, managers, and technical experts by surprise. Because occasional telework was so incremental a change to normal procedure, and required so little institutional adjustment, they assumed that going to a virtual, network-centric organization would be similarly trivial. Not so. Like any complex system, at some point one runs into nonlinear challenges, and that of course happened in this case as well. Figure 10.3 provides a sense of some of the adjustments that were required.

Consider that informal telework on a partial basis can usually be worked out between an employee and a manager, and implemented by carrying home a briefcase full of work and a laptop. But a network-centric organization requires that all the information an employee needs, from employee records to corporate policies to data and project information, be available 24/7, be secure, and be in useful format; that the human resources department knows enough about the employee to know where that person is located for tax and pay purposes; that the property management department knows which buildings are coming off long-term lease so they can be closed and operating costs shaved; and that the ICT department be able to keep the network up and running, and

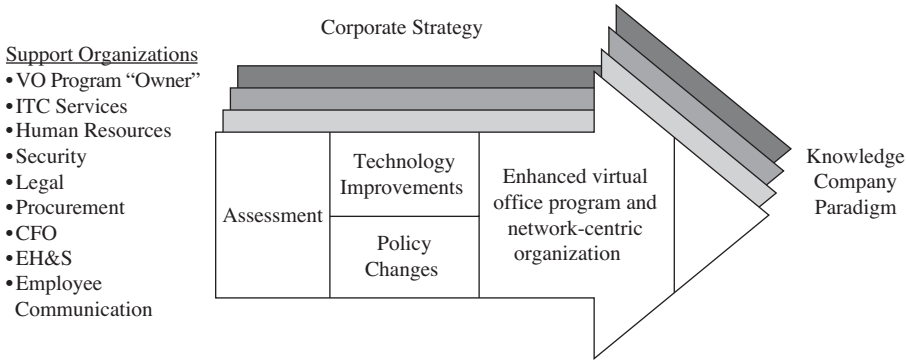


FIGURE 10.3 Multiple Dimensions of Business Transformation. Based on experience gained with virtualizing the workplace at AT&T in the late 1990s and early 2000s, this figure illustrates the process by which businesses are transformed from hierarchical place-based structures to network-centric virtual structures. Changes in technology are critical, but must be aligned with changes in corporate policy and corporate departments if meaningful gains in efficiency and quality of life are to be achieved. An important observation is that, at some point, the simple task of making it easier for people to work remotely triggers a transformation throughout the business; an incremental change becomes a shift between two qualitatively different states.

the corporate information technology help desk effective, under significantly more load. More subtle but perhaps more important, management practices and skills need to be upgraded; the lazy manager in most firms usually uses "time at desk" ("TAD") techniques to judge an employee's value, but a virtual manager needs to understand what the employee actually contributes to the firm. Most managers are, sadly, very poor at the latter.

Remember, however, that the more a desired change is made to look incremental rather than discontinuous, the easier it is to introduce. Thus, for example, in a firm unused to virtual office practices, the sustainable engineer might want to introduce new virtual office technology and systems as if that engineer were simply improving existing practices of informal work at nonoffice locations, rather than trying to jump to completely virtual systems all at once. Even though the sustainable engineer might understand the depth of change required, it is still more effective to package the change as incremental and in line with existing corporate culture and practices, rather than as something discontinuous. Managing such technology-led institutional changes, by the way, should not be considered "someone else's job": it is the job of the sustainable engineer to be not just technically competent, but a recognized leader able to manage such complex systems in the real world.

But, of course, for the sustainable engineer there is also an initial analytical issue here, and that is whether these sorts of practices are, indeed, more "sustainable," at least at first blush. The data in this instance, as always, are somewhat sketchy and ad hoc, but experience in a number of firms that have implemented sophisticated virtual workplace programs indicate that the practice

has significant **triple bottom line** benefits. In other words, working from alternative locations provides economic benefit to the company and the worker, social benefit to the community, and reduced environmental impact.[3]

Begin with a dimension that must be of concern to any engineer, the economic. Many companies, especially the early adopters, did so for financial reasons. IBM, for example, has since 1998 positioned over 15 percent of its workforce globally to work any time and any place, claiming almost $1 billion in savings per year, primarily in avoided real estate costs. Jet Blue Airways reduced the costs of booking a flight by 20% by using home-based reservation agents instead of facilities-based call centers, thus helping it to report a profit only six months after its first flight. Detailed studies at AT&T identified three major areas of economic benefits to the firm: reduced operating costs, especially in terms of real estate; increases in productivity, including less sick days taken by digital nomads, and greater per employee and per hour productivity; and enhanced personnel recruitment and retention. The economics of this program were quite favorable: real estate savings alone added up to about $35 million annually as buildings rendered surplus by virtual office initiatives were sold or came off long term leases.[4] Increases in productivity were worth $150 million per year and arose from two sources: teleworkers reported more productive hours per day, not surprising given that they saved substantial commute times that they divided between the firm and their personal activities. At AT&T, this reported difference was substantial. Teleworkers reported 7.7 productive hours per day, while office workers reported 6.8 productive hours per day. AT&T teleworkers also tended to be more productive on a per hour basis as well: 67 percent of all teleworkers reported higher productivity at home while only 3 percent reported higher productivity in the office. Of course, self-reported data are always somewhat questionable; more fundamentally, knowledge worker productivity is a difficult thing to measure. These productivity data were reinforced, however, by the managers of teleworkers, who also reported that teleworkers were more productive, and by personnel performance data, which showed that teleworkers at AT&T were higher ranked on performance than their office-bound peers. Moreover, these results were in line with more general

[3]This evaluation will draw heavily on data from AT&T, one of the first companies to rigorously implement a virtual office program for its managerial staff (organized employees chose to be excluded because of concerns regarding company controls extending into their private lives) and collect data on the effectiveness of virtual, part virtual, and digital nomad work practices. These data are not comprehensive, but are the best of which I'm aware. These data all predate the 2005 merger of AT&T and SBC that, as would be anticipated, created a discontinuity in the programs and data flows as human resource policies, staffing levels, real estate overlaps, and other issues were managed.

[4]In a "down" property market, with low demand for office space, it may be impossible for a company to sell a building it owns outright. On the other hand, a long-term lease is a contractual obligation; when the lease term runs out, if the firm can send those employees virtual, it does not need to renew the lease and can save those rental fees going forward.

findings from other companies pioneering virtual office programs, including Siemens, Compaq (now part of HP), Cisco, Merrill Lynch, Nortel, and American Express, all of which reported productivity increases of between 10 percent and 50 percent for telework programs. The economic implications of telework for the firm appear to be firmly positive.

Moreover, although more difficult to measure and highly dependent on the way programs are structured, the economic and certainly the quality-of-life implications for employees appear to be firmly positive as well. Most obvious, when offered a choice, many employees freely choose some form of virtual office arrangement. From the perspective of the firm, the benefits of enhanced quality of life and personal freedom for employees shows up as enhanced recruitment and retention. Again using AT&T data, one sees both necessity and choice: 19 percent of new teleworkers said they began the arrangement because their family situation changed and virtual work arrangements became essential, while 63 percent of teleworkers claimed higher job satisfaction as a result of virtual office programs. Moreover, from a sustainable engineering perspective one should also consider personal satisfaction and fulfillment: 67 percent of AT&T's virtual workers said that it made them feel more trusted, and 72 percent said they were much more satisfied with their personal and family life.

It is much more difficult to understand the social implications of virtual office initiatives, in part because they are so varied and touch such different subgroups within society. For example, a major and underappreciated benefit is the way that virtual office infrastructures and services enable society and individual firms to access knowledge workers (e.g., the disabled and retired populations) that previously were difficult to incorporate into economic activity. Remember that our discussion of demographics in Chapter 2 indicated how changing age distribution in developed countries was creating the potential for major financial dislocations as fewer and fewer younger workers had to support the pensions of more and more retired workers. An added cost of such an economy is that the institutional knowledge that is often implicit but critical if institutions are to continue operating, leaves with retiring workers, and cannot be replaced. Today, many firms and governmental entities around the world have a number of highly skilled and knowledgeable older workers approaching retirement age. If these resources become inaccessible, the impact could be devastating. The irony is that, in many cases, older citizens want to continue working, they just don't want to continue working in the same way they had to in the past. The interests of disabled potential employees are similar, as technologies evolve that help them participate fully in the knowledge economy, even if they are mobility, visually, or otherwise challenged. With younger workers, the essential desire for freedom from unnecessary temporal and spatial constraints, and to be able to balance life and work, is analogous, albeit for different reasons. Accordingly, the companies that are able to shift to virtual office infrastructures

and solutions will be able to differentially attract and keep good knowledge workers. These firms will also gain competitive advantage; at a cultural and national scale, a global economy that is continuing to become more knowledge intensive will mean that such a competitive advantage among firms will become comparative advantages between cultures and nations.

Not all social indicators are positive. An obvious issue to be managed is network security: The more reliant a firm or institution becomes on network functionality, the more vulnerable it is to attack. (In general, information networks that display a scale-free architecture are not affected by accident, unless it strikes one of their superhubs, a facility or piece of equipment that anchors substantial network functionality.) Many utilities, for example, have supervisory control and data acquisition (SCADA) systems that are informally or even formally attached to general corporate networks that are also attached to the greater Internet; this structure means that the systems controlling important parts of their infrastructure can be accessed if a hacker or enemy agent can get into the corporate network (which is usually not that hard if cyber warfare is involved). Where possible, important operational networks should be completely separate, both in physical and information flow domains, from Internet-connected networks; if that doesn't work, the connection should be explicit and managed (e.g., people must take memory sticks with the relevant information physically over to another machine); where that is not feasible, security systems designed for the relevant threats should be implemented.

As far as the final leg of the triple bottom line, the environmental implications of virtual office programs and services seem to be positive, at least in the short run. Most studies show that virtual office programs—which in practice tend to involve a significant component of working from home—does significantly reduce automobile travel, even given that digital natives obviously don't always stay at home. Again, using AT&T data circa 2003, its telework program saved 100 million miles of unnecessary driving, avoiding the consumption of over 5 million gallons of gasoline and eliminating the unnecessary emission of some 50,000 tons of carbon dioxide. Equivalent reductions in other pollutants, such as the nitrogen oxides and hydrocarbons that contribute to photochemical smog, also resulted. These data probably understate the affects, given that eliminating unnecessary commutes removes a car during the most congested part of the day. Not only is the commuter's car most inefficient in congestion and stop-and-go traffic, but congestion is a network phenomenon, so removing cars from peak congested periods improves the efficiency of the commute of every other car on the road. Shaving peak traffic loads also potentially reduces the need to build new transportation infrastructure, because roads, bridges, and the like are generally engineered for peak period capacity. Thus, virtual office programs can directly reduce infrastructure requirements over time.

This should give you, as an up-and-coming sustainable engineer, an important insight: Where such synergies exist across different technology and sustainability systems—in this case ICT services, traffic management, air quality, and transportation infrastructure planning, design and operation, for example—the engineer should learn to approach the systems holistically, so that you can optimize across them rationally and efficiently. Of course, in today's world we don't operate this way, and you can readily predict strong opposition from existing departments, governmental units, and professionals to the changes necessary to do so. This example is not just an indication of the opportunities for sustainable engineering, but also hints at the difficult institutional challenges to implementation of even fairly obvious sustainable engineering initiatives.

That virtual office/digital native services and practices produce integrated economic, environmental, and social benefit, yet are difficult to implement, should be no surprise by this point. Any fundamental evolution—and virtualization of the workforce is a major part of the shift from an industrial to a knowledge-based economic structure—takes time and will be resisted by existing cultures and institutions. Moreover, as with any complex change, virtualization of work does not advance by itself; rather, it requires an integrated evolution of practice, technology, culture, and institutional structures, not unlike the core technologies that power Kondratiev waves. Where systems are thus tightly linked to each other, evolution of necessity is a slower process than if the links didn't exist.

10.6 VIRTUALIZATION OF WORK AND THE EVOLUTION OF INSTITUTIONAL COMPLEXITY

So far, we have followed changes in ICT technology platforms, to new services enabled by those technologies, through a sustainability analysis that provides some idea of the complexity of these integrated technology systems (e.g., the linkage between virtualization of work, life extension, and continued economic productivity), to institutional change. This is already a rich case study for those interested in sustainable engineering. But even this only scratches the surface. To understand virtual office implications at the level of the firm, it is first necessary to reflect on some of the fundamental trends currently shaping firm structure and organization, because virtualization is both a necessary adaptation to these trends and a causative agent of structure change in its own right. The two most important trends are the substantially enhanced complexity of the context within which the firm must operate and a concomitant evolution in firm structure from highly bounded, facility-based entities where physical assets by and large determine value, to a rapidly evolving, information-based network-centric structure.

Have you ever stopped to think about what a company actually is? We tend to think of a company as a thing, but in many ways it is a pattern, one that changes as the environment within which it operates changes (and, confusingly,

one that it helps to shape over time). Firms can thus be thought of as arising from and enabling particular levels of complexity in economic, cultural, and governance systems. In the early days of the Industrial Revolution, for example, firms in Europe were essentially factories (hence the name "factory capitalism" for this period of economic development). To the extent that larger operations were required, one-off institutions were created, such as the royally chartered British [or Dutch] East India Company. These were essentially extensions of the State: slow to adapt, clumsy in management structure, and militarized. But, in part because of the wealth and economic activity generated by these early forms of the firm, subsequent growth in complexity led to the need for more and different patterns, more adaptable and not defined directly by the State. Thus, because economies kept getting more complicated, and technologies kept getting larger and more complex in their social and institutional implications (remember the railroad), it became necessary to evolve general incorporation statutes. In the United States, the first such statute was passed in New York State in 1811. The importance of this from the perspective of complex adaptive systems cannot be overstated. Without such generic firms, competing with each other and thus creating conditions for institutional evolution, we would not have the economy we have today, nor would we be able to support the population, the technology structures, and the quality of life that we have today. Just as the form of money evolved to enable and reflect greater complexity (Box 5.3), so the form of private firms evolved for the same reasons.

But as the economic, technological, and social context continues to grow in complexity (as discussed in Chapter 3), it is no surprise that the existing pattern of the firm—hierarchical, relatively slow and resistant to institutional change, not very agile in exploiting new markets and technologies—is increasingly undermined by the very forces that those firms have created. The old model of the firm and its relationships, defined by hard boundaries between different economic entities, between producer and consumer, between workplace and anywhere else, between employee and nonemployee, was well suited up to a certain level of complexity. In systems terms, it provided local stability—in time, space, and information dimensions—that permitted the evolution of a highly complex globalized economy predicated on manufacturing and artifacts as primary sources of value. But it is now obsolete in many ways (one example of this for those interested in industrial organization is the failure of the tightly controlled model of a firm that characterized Route 128 around Boston to compete with the much more flexible structures that characterized Silicon Valley). And so, one moves to the question of sustainably engineering the institution of the firm for a far more complex environment that is itself a product of previous archetypes of the firm.

Figure 10.4 illustrates some of the implications of the network-centric firm. Remember, however, that the network-centric firm is an ideal type, a pattern that

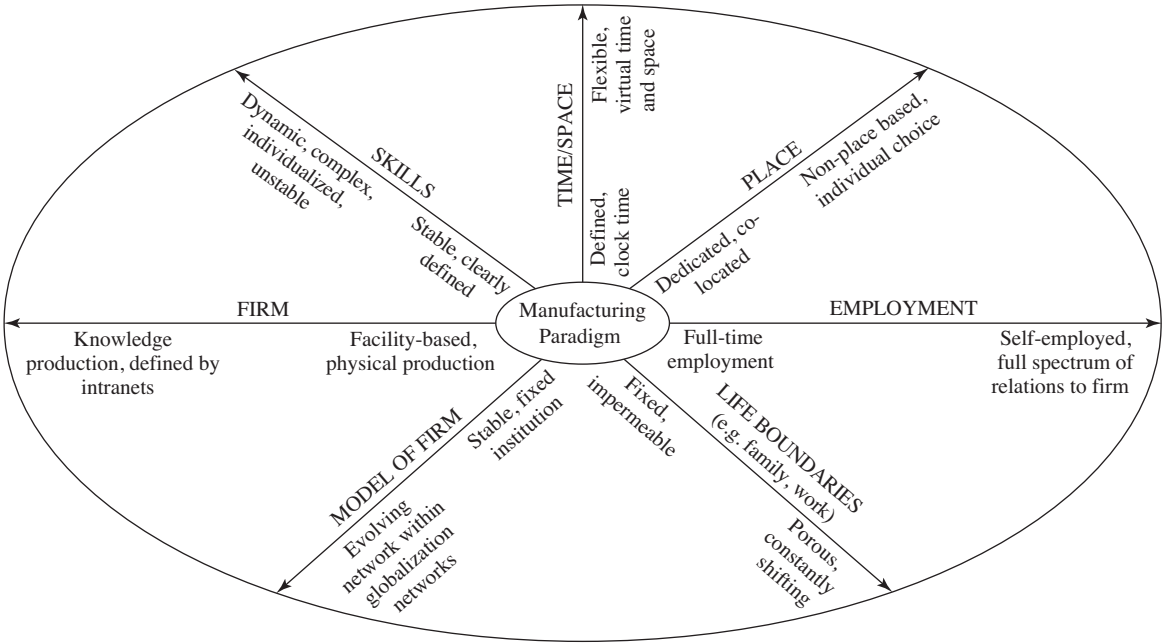


FIGURE 10.4 Virtual Office Social and Institutional Implications. This figure illustrates some of the dynamics implicit in the evolution towards the virtual workplace. This example is chosen in part because the underlying change—basically, work at home—is relatively intuitive to many people, yet the institutional implications of that shift are profound and almost never considered, even by people familiar with the technologies and policies involved.

may apply to parts of a firm's operations and not others. Obviously, for example, a manufacturing firm might rely on virtualization of work to reduce demand for pure office space and move its supply chain management operations to Web-based systems, but it will still operate factories somewhere. Moreover, given the competitive advantages of virtualization, such as agility in the face of changing market conditions and lower operating costs, balanced against the observation that no institution changes unless it has to, one would anticipate that the most competitive parts of the firm's operations would change first. Finally, like any complex system a firm will thrive with a somewhat hybrid structure: some virtual and nonhierarchical components, and some hierarchy providing structure and support. Given these caveats, what are some of the implications of the superficially simple shift from place-based to virtual workplace?

Beginning with the conceptual model of the firm itself, the obvious shift is from a fixed, defined institution that is stable through time along a number of dimensions (e.g., lifetime employment, community-defining manufacturing facilities) to a network that shifts in response to changes in internal and external state. Note that this requires a fundamental shift in mental models. The model that brings white-collar knowledge workers in to their desk every morning in the same building is basically the manufacturing model. In manufacturing, of course, if you're not working the line, you're not being productive. But this is a different economy, with a different production model, and that manufacturing mental model is not just obsolete, but actually harmful. It leads to people wasting their time in unnecessary commutes, burning unnecessary gas, or taking up unnecessary space on busses. It also leads to dissatisfaction, because employees are being judged on manufacturing metrics ("How much time did you spend at work?") that are valid for the manufacturing environment, but not for the knowledge economy (Do you stop thinking creatively as soon as you walk out of your corporate office? Do you stop thinking as soon as you walk out of the professor's lecture?[5])

Firms used to be defined by their physical property such as buildings and inventory, because that represented their economic value; firms now increasingly stand or fall based on their ability to use nonphysical assets, such as patterns of information or skill in managing information. Corporate efficiency increasingly relies on productive use of knowledge and an internal structure defined more by corporate intranets than by facilities or hierarchical organization charts. Walmart, for example, is the world's leading retailer right now not because of its savvy placement of "big-box" retail outlets, but because it operates a better ICT operation across its global supply networks than any of its

[5]We will avoid the delicate question of whether you need to think during the lecture. The need to re-engineer education to reflect the cognitive changes ICT has created is a real one, but too large to address in passing in a sustainable engineering text.

competitors. More generally, efficient firms are shifting functionality to corporate intranets, with a concomitant savings in time, accessibility, and even paper. Network-centric firms, new models of working, different ways of managing corporate information—indeed, new models of corporations that lack boundaries entirely—this is corporate structure in a knowledge economy.

This is a more complex world for the individual as well. In the old manufacturing model, work was defined and specific. After all, Adam Smith and many subsequent economists noted that an important initial driver of efficiency in capitalism was division of labor—giving each worker a specific, repetitive, bounded task. But now, the employee herself becomes more of a node in a complex network. Skill sets are no longer stable and clearly defined; rather, they also become dynamic, individualized, and subject to change as requirements of the external and internal subsystems change. The world is changing faster and faster: If you want to stay relevant, you have to as well.

This leads to the observation that a network-centric structure leads to a firm that no longer operates on clock time at a dedicated place, but rather is neither place based nor time based. In some ways, this simply reflects globalization, because transnationals operating around the world have no "home" time or place; rather, their work migrates around the world as different units in Asia, Europe, Africa, and North and South America come on line during the 24-hour day. But more broadly it reflects the shift from an industrial to a knowledge economy: cyberspace is open 24/7, and those that don't learn to operate within it eventually fall to those who do. Again, compare the manufacturing and knowledge economy models: Manufacturing requires colocation in time and space of those employees who work in a manufacturing facility and create products. If you are not on the assembly line when required, you are not productive. On the other hand, a knowledge economy benefits from the intellectual capital of the knowledge worker —and that is neither time nor place specific. Good ideas may come in the shower, or on vacation in another continent, or while falling asleep on a flight home from an assignment, and it can come from someone in Russia, China, Brazil, Indonesia, or Singapore. It is the ideas, the intellectual capital of the knowledge worker, which the network-centric firm must capture. This broadening of the firm is not unconnected to a realization that the old model of customer distribution—sell to those in developed countries who can afford your product—also requires modification. If one is smart and adaptive, one can sell many products, suitably designed, not just to the well off, but to those "at the bottom of the pyramid" as well (see Box 10.4).

Finally, it is apparent that the network-centric firm implies the breakdown of the previously strong boundary between work and individual life for many people. When one works on a manufacturing line, it is quite apparent when one is at work and when one is not. But the virtual workplace has no

BOX 10.4 ICT at the Bottom of the Pyramid

In 1932, President Franklin D. Roosevelt in a speech referred to the need to consider those at the "bottom of the economic pyramid," a phrase taken up in 1988 by Stuart Hart and C. K. Prahalad when they focused attention on the economic opportunities offered by the world's poor, who were generally being ignored by firms more used to middle- and upper-class customers in developed countries (hence the name of Prahalad's 2004 book, *The Fortune at the Bottom of the Pyramid*). ICT growth in developed countries offers a view of what this means for sustainable engineering.

We can begin with the economics, which are vastly different. In many developed countries, the monthly revenue from a mobile phone subscriber is between thirty and sixty dollars; in developing countries, the revenue is between three and fifteen dollars (figures circa 2010). Subscriber growth, however, has straightlined in developed economies, but continues to grow robustly in developing countries. From a social perspective, adding even small amounts of ICT capability dramatically increases not just personal economic advantage, but national GDP; some estimates indicate that adding one telephone per every ten people in a developing country boosts GDP growth in that country by 0.6 percent to 0.8 percent. Similarly, some studies have shown that access to mobile phones on the part of fishermen in developing countries, by enabling more efficient sales of their catch, reduce consumer prices by 4 percent and increase profits for the fishermen themselves by 8 percent. Firms providing the service, of course, also benefit, albeit they obviously require different business models and mixes of services.

So what does operating at the base of the pyramid require? First, it requires technological prerequisites, in particular standardized, inexpensive technology platforms (which may, in turn, require evolution of technology standards by industry). Second, it requires changes in business models. For example, in many developed countries, the standard billing model for many services and products involves a form of extended credit; charges are "postpaid," that is, you pay your bill after you have used the service, usually on a periodic basis (e.g., monthly). But this is not practical in developing countries, which operate so-called "cash economies" instead of "credit economies," so "prepaid" billing systems needed to be introduced. This enables poorer people to purchase service as they need it, and avoids the enormous infrastructure required by billing, debt collection, and other "back-office" functions necessitated under a postpaid system.

Second, it may require deep innovation. In developed countries, mobile phones are evolving from telephones without wires into sophisticated information appliances, with many uses involving entertainment, information overlays, music and video downloads, and other services. In developing countries, however, the nascent information network implied by mobile phone technology has far different, and more serious, services platformed on it: money and banking services for those who would otherwise have no such options; important health care information; and increasingly, political and social information that empowers populations

that are otherwise easily controlled and cowed. The technology is the same (only much less expensive); the services dimension profoundly different; and the social consequences of the technology thus also profoundly different.

And what should the sustainable engineer look out for when working on such innovative products and services? To begin with, it is always a general rule that incumbent technologies and institutions will oppose new competitive innovation regardless of the eventual benefits. Thus, for example, among the strongest opponents of money services via mobile phone technology are banks. And what can be done about that? Depending on the situation, one can go head-to-head with entrenched interests—always risky, especially if the entrenched interests are coupled to governing elites—or instead, co-opt them. Thus, for example, a new mobile phone company seeking to offer financial services at a micro level via mobile phone might partner with a powerful local bank, turning an opponent into an advocate.

More subtly, it is frequently very difficult for innovation to arise in institutions that are deeply immersed in, and successful using, existing technologies and business practices. In regard to mobile phones, the most successful operators in Africa and Asia are generally not the European or American giants, but indigenous firms with a solid knowledge of the local customer base and political intricacies. Such firms, nimble and keyed to their customers, can perceive and enter markets to which others are blind: Local mobile phone companies have prospered in war-torn regions of Africa such as the Congo where civic order has long been absent, and no larger foreign firm would dare enter. Moreover, it may also be necessary to play with business models. Many developed country mobile phone operators, for example, own all of their own facilities, a business model drawn from the landline companies, and profitable in economies where each subscriber may be paying fifty dollars a month. But for a developing country operator, controlling costs is an absolute necessity, so all possible operations are outsourced, including in some cases even ownership and operation of the physical plant.

The bottom line for sustainable engineering? There is an enormous amount of social, economic and, indeed, environmental benefit to be generated by operating at the base of the pyramid, and for-profit operations can play a critical role in achieving these benefits. But difficult institutional challenges need to be understood and overcome, local cultures and conditions respected, and many assumptions about price/performance levels, behavior, institutional stability, pricing, marketing, and other aspects of engineering questioned in making the transition to such markets. In addition, as the mobile banking services indicate, innovation is as important in developing markets as in developed markets, and it requires the same attention not just to technology, but to institutional, social, and market context as it does anywhere. Developing markets, in other words, are the same as developed markets, only different.[6]

[6] I meant to say that. Cognitive dissonance is not always a bad thing; sometimes it is the essence of insight.

such boundaries: work time is when the computer is on and email is getting answered, or when relevant Facebook posting are being done from an airport in Europe or Asia. Neither local place nor local time defines a status as worker or nonworker anymore.

This collapse of the boundary between work and family is not new. Indeed, it is quite old. A physical and temporal integration of work and family characterized much of human life in hunter–gatherer societies, in agrarian societies before the Industrial Revolution, and the cottage industry pattern of work in Europe just prior to the Industrial Revolution. Nonetheless, the implications are difficult, and the adjustment is more complicated than it appears. For example, it is a strength of virtual workplaces that they enable individuals who may be disabled in a number of ways—mobility or visually challenged, for example, or seniors who may not be as active as before—to participate without discrimination in the workforce. Yet many disability laws require that individuals on disability benefits not work, even if the disability is physical and the work is mental. And taxes on income are all place based, so a mobile worker becomes an immediate headache for the firm trying to direct taxes on income to the appropriate jurisdiction.

This is not just a technical taxation problem, but a difficult conceptual one that is somewhat analogous to the problems that arise when national boundaries and laws clash with global virtual realities. Network-centric firms no longer have a boundary they can rely on to define their legal responsibilities. If a firm is defined by a manufacturing facility, the geographic scope of its liabilities is clear and defined by the boundaries of its property. If you get hurt working at the factory, one set of laws (workers compensation in the United States) applied; if you get hurt at your house, another. But if a knowledge worker may be performing company business anywhere, from a cybercafe when on trek in Latin America, to a coffee shop in California prior to a surfing afternoon, or from the airport returning from a vacation—just where does the firm start and end? That question has been answered in a manufacturing economy, and the answer embedded in society, law, and behavior; it has yet to be adequately asked for the knowledge economy and virtual realities and spaces.

Finally, consider some of the technology challenges. Although these are numerous and constantly changing, there are several that reappear on a regular basis. Among these are bandwidth, security, and maintenance of network functionality. In the past, lack of broadband was the top barrier to increased telework cited by employees. This has receded to a second-tier issue in many developed countries, but it remains important because the principle assumption behind true virtualization of work (or any activity) is that functionality should be independent of location. Moreover, there is never enough bandwidth, because function expands in unpredictable ways as bandwidth

increases. Facebook and other social networking sites had to wait until broadband was available to a critical mass; who knows what interesting functionality will be developed when the network can provide full interactive, real-time video (or holographic streams, eventually) on demand? The importance of network security has already been mentioned. And, of course, the more an institution is dependent on a network, the more important it is to keep the network up and running, and the individual's connectivity operational—and the harder it is, because one is moving from the relatively closed structure of a building or firmwide network to the much more open environment of connectivity everywhere.

10.7 CONCLUSION

This chapter has only begun to scratch the surface of how ICT as a system, from material to artifact to network to service to shaper of culture, is changing the anthropogenic Earth. It does not, for example, discuss the impact of the Internet on old business patterns, and the efficiencies and issues associated with business to consumer (B2C) and business to business (B2B) commerce. Moreover, especially the last section may raise a question for the student: What has this got to do with sustainable engineering?

The answer is very simple. This is a complex world, and although there are many times you will be operating in one particular bounded patch—say, designing a new mobile personal assistant—your work will be connected in many ways to many impacts across a much larger system. It is often the case that those connections can be safely ignored, but it is increasingly the case that they cannot be, either for practical design reasons or because stakeholders won't let you. For example, when the author of this text was in charge of introducing virtual office practices at AT&T, it was not just a matter of finding the right technology; it was necessary to come up to speed very quickly on institutional change, because what was initially taken for a relatively simple change in work patterns turned out to be much more profound. If you launch the successor to Facebook, you had better understand not just your technology and your customers, but also how to ride tides of cultural change that will either mark the next social networking wave or leave you stranded in the past as the future roars by you. And don't kid yourself: such products change the world.

More important, the sustainable engineer in an age of unpredictable, rapid, technological change cannot just be the closet expert, locked into a team deep in the organizational chart and tasked to find the right bolt for the final design. You have a responsibility to yourself and to your society to be an ethical, responsible, and rational leader in a world increasingly shaped by the technologies that you and your peers create. And to do that, you can't stop with the artifact, you must strive to understand the system. That means being open to

learning and change as part of your professional expertise. It means understanding not just the networked technology platform that powers the network-centric organization, but the levels of institutional and personal change that surround it. And it means understanding that, as an engineer, you are not simply a designer of things; you are, in a profoundly creative act, bringing into being one future out of many possible futures. Individually and collectively, we can no longer refuse to take responsibility for the world we are building. That's the sustainable engineer as leader.

Exercises

10.1 You are the sustainable engineer for a firm that manufactures laptop computers in the United States. Your CEO asks you for an analysis of the design of the company's products. What would you establish as the boundaries for your analysis, and why?

10.2 You have finished the analysis your CEO asked for in Exercise 10.1, only to find that the manufacturing operation has been sold to a Chinese firm, although the U.S. firm will continue to brand and sell the laptops. How might this change the boundaries of your analysis?

10.3 Still working for the same company as in Exercise 10.1, you find that a small yet critical component for capacitors in your laptops is derived from tantalum that is sourced primarily from Central Africa. The profits from mining it help finance armed militia who routinely engage in murder, torture and rape in the course of conflict. Should this be part of your sustainability analysis? Why or why not?

10.4 You are the sustainable engineer for the firm that supports a highly popular massive multiplayer online game (known as MMOG in the trade), *Orcs and Oracles*. Every day well over a hundred thousand players inhabit your world, which is housed on server farms located in the American Southwest and sub-Saharan Africa. Create a map similar to that shown in Box 10.1 that illustrates your sustainability effects.

Annotated Bibliography

1. The best source on technical aspects of DFE is the IEEE, 1993–2008, *Proceedings of the Annual International Symposia on Electronics and the Environment*. From 2009 on, the conference from which the papers have been drawn has been renamed to capture the increasing importance of sustainable engineering; the successor publication is the *Proceedings of the Annual International Symposia on Sustainable Systems and Technology*.

For those of a historical bent, perhaps the first glimmers of DFE were the American Electronics Association White Papers circa 1991: "What is Design for Environment?" (B. Allenby, AT&T); "DFE and Pollution Prevention" (B. Allenby, AT&T); "Design for Disassembly and Recyclability" (R. Grossman, IBM); "Design for Environmentally Sound Processing" (J. Sekutowski, AT&T); "Design for Materials Recyclability" (W. Rosenberg, COMPAQ, and B. Terry, Pitney-Bowes); "Cultural and Organizational Issues Related to DFE" (B. Allenby, AT&T); "Design for Maintainability" (E. Morehouse, Jr., USAF); "Design for Environmentally Responsible Packaging" (K. Rasmussen, General Electric); "Design for Refurbishment" (J. Azar, Xerox); and "Sustainable Development, Industrial Ecology, and Design for Environment" (B. Allenby, AT&T). They were subsequently published as a collection by the American Electronics Association in 1993. These were initial efforts and relatively simple by today's standards, so they are more of historical interest than anything a reader would actually want to rely now.

2. There are a number of books dealing with the information economy, knowledge economy, or whatever you prefer to call it. A good introductory book is C. Shapiro and H. Varian, *Information Rules: A Strategic Guide to the Network Economy* (Cambridge: Harvard Business School Press), published in 1999; a broader treatment and an excellent introduction to the networked economy is M. Castells's 2000 book, *The Rise of the Network Society* (2nd ed.) (Oxford: Blackwell Publishers).
3. There is a specialized literature on the virtualization of work; among those that may be of interest are M. Igbaria and M. Tan's somewhat dated but still useful 1998 *The Virtual Workplace* (London: Idea Group Publishing), and on the legal issues around telework, N. Goluboff's 2001 book, *The Law of Telecommuting,* published by the American Bar Association. Two other volumes that may be of interest are *Information Technology in the Service Society* by the U.S. National Research Council in 1994, and D. Richards, B. Allenby, and W. D. Compton, *Information Systems and the Environment*, written for the U.S. National Academy of Engineering in 2001; both volumes were published by the National Academy Press.

CHAPTER

11 The Five Horsemen, Military Operations, and National Security[1]

You may not be interested in war, but war is interested in you.

—Attributed to Leon Trotksy

11.1 INTRODUCTION

The last several chapters have dealt with two technology systems, ICT and green chemistry, which are implicitly framed in terms of the core technologies characterizing each: entertainment and information production, and computer and communication technologies for the former, and chemical engineering, pharmaceutical design, and nanotechnology for the latter. This chapter shifts gears, and frames its discussion in terms of intended functionality: in this case, use in military operations and in the broader venue of national security. It thus introduces not only a new substantive area for consideration, but also a different way to conceptualize

[1]This chapter, and some of its graphics, draw on work done by the author for the Consortium for Emerging Technologies, Military Operations, and National Security (as does, self-evidently, the title). In addition, parts of this chapter are based on essays written for and published on the Arizona State University Consortium for Science and Policy Outcomes Web site, and on a chapter in another book, *The Techno-Human Experience*, written by B. Allenby and D. Sarewitz.

technology systems. Which one is most useful depends on the nature of the project that the sustainable engineer is working on, and why that engineer is working on that project. (Construct something new? Add to the resilience of an urban infrastructure or system? Reduce the chances of a terrorist attack, or the impact if such an attack does occur? Reduce battlefield casualties? Deter potential aggressors? Be successful as a private military services provider?)

Let us begin, however, with an obvious question: Does a chapter on military operations and national security belong in a text on sustainable engineering? Certainly, many sustainability advocates would answer in the negative. If the sustainability community is apathetic regarding technology generally, it is usually hostile to military institutions and operations. So far as that goes, it is also true that many textbooks on engineering tend not to discuss military technologies and related issues. In both cases, however, this is both shortsighted and unsystematic. Engineering in civil society and military engineering have over history traded technological leadership. In the 1500s and 1600s, for example, European civilian ship design fed back into military applications (in fact, in the form of privateers, the two were sometimes the same), but by 1900, with a full-blown naval arms race between Germany and the United Kingdom, naval systems engineering was driving the evolution of civilian metallurgy and engineering. Skill in casting church bells in medieval Europe was a crucial technology enabling the subsequent casting of cannon; later, the naval power race (in this case, primarily between France and the United Kingdom) would be an important source of new material technology.

In most ages, in fact, military and security issues are critical drivers of technological evolution. This is especially true in our world, where weapons of mass destruction have been commoditized and cultural conflict is endemic. In such an environment, as some of the examples discussed below illustrate, military and security technologies may have far more to do with sustainability of institutions, macroethical challenges, individual and human rights, quality of life, and cultural evolution, than more popular obsessions such as climate change.

At the most basic level, military activities lead to technological evolution and dominance of particular cultures over others with impacts that strongly affect subsequent history. An obvious example is the development of the oceangoing caravel in the fifteenth century by the Portuguese, which enabled open ocean travel by Europeans and, equipped with cannon and gunpowder (especially once "corned" gunpowder, a much more effective formulation, was introduced), was an important mechanism for projecting force and thus enabled trade empires and the economic basis of imperialism. Development of the chariot, combined with evolution of the bow from simple to composite—a development in bow design that resulted in powerful bows small enough to be used from chariots or horseback—dramatically changed power relationships. When the Greek city-states developed military strategies based on hoplites and

hoplite technology—citizen infantry that stood shield to shield, usually armed with spears and secondarily with swords—they not only defended their city-states, but made a profound statement about their culture. The charioteer represented what historians term "chariot nobility," and the infantry supporting him was composed of subjects of the ruler, but the hoplite, a citizen soldier defending his home city, represented the arming and the elevation of the subject to a citizen. A chariot war was a war of the nobility, and thus supported such governance structures; a hoplite war was a war of the citizens, and thus implied either a communistic (e.g., Sparta) or democratic (e.g., Athens) polity.

In our own time, perhaps few military technologies have had the impact of the nuclear bomb. In its initial use, it killed tens of thousands (although conventional bombing of cities had actually killed more in some instances). The decision to use the bomb was in large part a strategy to avoid a ground invasion of Japan, with potentially huge loss of life, a decision over which people are still arguing. But it was far more than that: It was the moment when the reality of the anthropogenic Earth became evident to everyone. Earlier in 1945, when the Trinity Test of the first atomic bomb was held in the deserts of New Mexico, Robert Oppenheimer, one of its designers, quoted the god Vishnu from the Hindu scriptures, the *Bhagavad Gita,* "Now I am become Death, destroyer of worlds." Note the profound shift in perspective in the two contexts: from Vishnu, a deity describing his godly power, to a mere mortal in awe not of what God or Nature had visited upon us, but what we had built for ourselves. And that technology not only changed military history, but the psychology of humanity itself. Now it was indeed possible for us to destroy not just each other on a group basis, as war had always been fought and as the bombing of Dresden, Hamburg, Tokyo, and London had all too clearly illustrated, but the world as we know it. A nuclear war and the ensuing nuclear winter would not just be a military event, but would be a cataclysmic event on a global scale.

Nuclear weaponry is thus a familiar example of the highly serious, complex, and rapidly evolving interplay among emerging technologies, military operations, and national security. The existential challenge to society represented by warfare, combined with the immediate advantage that new technologies can deliver, tends to accelerate technological innovation and diffusion. This is counteracted, in very interesting and educational ways, by the deep conservatism of military institutions and personnel, a conservatism justified by the high cost of any unsuccessful innovation in a military context (that is, you get it wrong, you die). Moreover, technology in a military context is not neutral regarding the decision to deploy force, either. If you think you are technologically superior to your opposition, you may well initiate conflict more readily. Many historians, for example, think that one reason the Bush administration initiated the conflict in Iraq was because it was convinced that its military technology would be so dominant that any opposition would melt away. In terms of

conventional military conflict, the Bush administration was right; "shock and awe" as a technology-based military strategy was amazingly successful. But technological dominance is much less effective in counterinsurgency situations, which the Iraq conflict rapidly, and predictably, became, leading to a long, expensive, and ineffective slog by the Americans.

In addition, the relationships between the military and security technology systems, and consequent social and ethical issues and changes, are quite complex. This is particularly true when, as now, technological change is both rapid and accelerating, posing the risk of cultural backlashes that could affect both short-term mission capabilities and longer-term security interests. A culture that rejects technological evolution is quite likely over time to find itself less competent against cultures that continue to innovate. Moreover, emerging technologies are likely to have similar destabilizing effects within the military, potentially affecting not just military operations, but military culture and organization, as well as broader social perspectives on military initiatives generally. This may seem a small concern, but because the military culture and institutions are the carriers of the ethics and rules that govern war, from following international agreements such as the Geneva Convention to the powerful idea that officers are personally responsible for what happens under their leadership, it can have a dramatic effect on the conduct of warfare, especially the impact of combat on civilian populations.

Finally, as has been illustrated earlier, emerging technologies are an important domain for sustainable engineering. If, as is the case these days, military needs are driving important aspects of the evolution of these technologies, the sustainable engineer cannot afford to be unaware of developments in military and security technology, and their potential impacts on civilian society and technology choices. This is especially the case where the professionalization of military activity has further separated military and civilian worlds, so that the informal awareness of military activities and culture that used to be assured by the draft and massive civilian participation in the military can no longer be assumed.

11.2 THINKING ABOUT NATIONAL SECURITY

Because this is a very complex area, let's begin by making a few observations regarding the relationship between military operations, usually immediate and quite explicit, and national security taken in its broad sense. This is important to our discussion because if a particular technology is helpful and ethical in an immediate military application (Level I or even Level II), but undermines national security in another way when viewed more broadly (Level III), a balancing of costs and benefits at various levels is necessary if a reasoned decision about deploying such a technology is to be made. And military success or failure does not equate to strategic success or failure. For example, many people

would agree with the claim that Japan and Germany "lost" World War II—as indeed they did, militarily. But both those countries are now wealthy, with democratic governments and citizens who are among the world's most productive and best educated, and are both strong pillars of the West. Similarly, few doubt that the French, and then the Americans, clearly "lost" Vietnam, yet that country is now an increasingly important trading partner of the United States, and if not democratic, certainly poses no challenge to American power. What "hard power"—military combat—appeared to lose, "soft power"—trade and market economy forces, and cultural attractiveness—achieved. There are numerous examples of powerful military states that have overextended their commitments because they focused on their hard military capabilities, and in the process overlooked the economic weakness that also resulted from that overextension, eventually bringing them down. Spain in the 1500s is one example; the United States in the early twenty-first century may be another. History does not judge "win" or "lose" nearly as quickly as combat does, and a systemic view of security must be careful not to jump to conclusions. A technology that enables effective Level I responses on the battlefield, but undercuts Level III national security issues, perhaps by encouraging a nation to overextend itself in global conflict, may be more problematic than one moderately effective on all levels. Such difficult analytical challenges make sustainable engineering in the context of military technologies perhaps even more difficult than in other areas. Box 11.1 explores the question of hard and soft power in more detail; for the student of sustainable engineering, however, this domain is of interest precisely because it is so complex, with so many different interdependent elements becoming dangerously contingent at once. Powerful technologies, an important province of sustainable engineering, often have security implications even where they don't have obvious military implications (remember railroads and the Prussian military). More important from the perspective of sustainable engineering, military experiences such as the Iraq conflict strongly warn against focusing too much on the technology, and not enough on the cultural and social context within which that technology is deployed.

One of the basic reasons that understanding emerging technologies and military activities is so difficult is that the very nature of warfare has changed. The sorts of wars developed countries fight these days are not traditional military-to-military conflicts in the classic European sense (Afghanistan and Iraq, Bosnia, the international intervention in Somalia, British and French intervention in African conflicts, and Vietnam are very different than the Napoleonic Wars or World Wars I and II). Unlike the constant resource conflicts in Africa between tribes, countries, and private militaries, for example, recent European and American conflicts have not been for conquest of territory or establishment of colonies, but to achieve ideological aims and defeat disfavored elites.

BOX 11.1 Hard and Soft Power in the Modern World

The general public perception that national power is predominantly a matter of military capability—so-called "hard power"—has long been recognized as oversimplistic and naïve by political analysts, who have increasingly focused on the importance of "soft power"—effective policies and institutions, and attractive cultural attributes—the power of the brand as opposed to the power of the gun, if you will. More specifically, five foundational sources of modern hegemonic power can be identified: economic, science and technology capability, military, institutional, and cultural. (There are other ways of parsing state power, of course, but this schematic breakdown works as well as any for our purposes.) Of these, the military component is "hard power"; the others are all "soft power." All five play important roles. Economic power linked to manufacturing prowess, for example, propelled Japan into the first rank of powers despite its relative military weakness, while political and financial institutional weaknesses have contributed significantly to its drift over the past decades. The European Union retains its dominant international position despite a relatively unimpressive military in part because its institutions reflect a resilient and open social-democratic political framework and its culture is widely admired (Europe has a good "brand image"). China's rise reflects an economic boom and extraordinary innovative energy, and those who predict a long lag until this is reflected in superpower status do so in large part because of perceived institutional weaknesses and because its brand image is relatively poor (few people given a choice opt for a repressive society). Even in the case of past empires, flexible governance structures, and the technologies and institutions that knit together distances, were as important as sheer military power. Roman roads and administration, for example, did as much to support the empire as did the legions, and the relative peace and prosperity of being part of the Roman empire in a chaotic and violent world made it highly desirable. England's domination of the globe from 1815 to the early 1900s owned as much to sophisticated financial and organizational innovations as it did to the English navy. That said, neither would have lasted long without the underlying military muscle of the Legions and the Navy, respectively, which created the environment within which civil institutions could safely operate.[2]

In a world where the most advanced economies are characterized by increasing reliance on information networks, highly flexible economic and political institutions tending towards the flat and virtual rather than the hierarchical, and

[2]It was, of course, on June 18, 1815, that the British under the Duke of Wellington held off Napoleon Bonaparte's forces until von Blucher's Prussian army joined them and completed the defeat of the French at Waterloo, thus ending the Napoleonic Wars. Perhaps the real key to Napoleon's defeat, however, was the battle of Trafalgar on October 21, 1805, where the British under Admiral Nelson defeated the combined French and Spanish fleets, thus assuring continued British domination of the seas and confining Napoleon to the continent. Not for nothing does the British Navy on every October 21 raise a toast "to the immortal memory."

reduced dependence on direct control of resources, the key to obtaining and keeping hegemonic power is balance among the five core constituents. This is arguably the key to the dominance of the United States, which until recently has been the one power that has appeared to be globally competent in all five categories: the largest single economy, a currently unmatched science and technology capability (and underlying academic infrastructure), a military that is far more advanced than any other in many ways, an institutional structure that is both relatively transparent and defined by law rather than relationship, and a cultural ascendancy that is reflected in the widespread idea of American exceptionalism, entrepreneurism, and corporate trademark dominance (especially in consumer culture: Coca-Cola, McDonald's, Disney). Although different actors focus on different aspects of this hegemonic power structure—the French, for example, who do not fear the United States militarily because it is an ally, tend to emphasize American cultural domination, while the Chinese, who face projection of American power and particularly its blue water navy as they look towards Taiwan, are much more concerned with its military aspect—it is the balance of high competence in all five dimensions that makes America relatively formidable.

But the essence of American power is more subtle than just the components in themselves, as the Japanese, Chinese, and European Union experience reveals. It is not enough to develop a better, more aggressive economy, as Japan in some ways did and China is doing; or to challenge American exceptionalism and cultural appeal, as Europe increasingly does. To truly challenge America over time, it is necessary for the rising power to become competent in all five dimensions. To complicate matters, all five dimensions are not independent of each other; rather, success in each requires synergism among all. A challenger, then, need not reflect the same balance nor certainly the same choices that characterize the United States, but it does need to successfully compete globally in all five dimensions—and to integrate them effectively so they are mutually supporting as well. This is a difficult challenge, even now when economic power is arguably shifting eastwards from America to Asia, especially China, and south to Brazil and perhaps someday Mexico.

To take an example relevant to sustainable engineering, consider the daunting task faced by a power trying to match the United States in science and technology (S&T). The American research and development budget, especially when expenditures from private industry are included, dominates global research and development (R&D) expenditures and ensures American supremacy in this critical area. But, worse yet from the perspective of those who would challenge America's S&T capabilities, it is buttressed by institutional and cultural dimensions. The American higher educational system pulls in intelligent students and graduates from other societies across the globe, many of whom remain in the United States because of its relative individual freedom and encouragement of innovation. The American venture capital system, again the most highly developed in the world, supports this structure by ensuring that S&T advances are rapidly translated into entrepreneurial activity and thus economic power. The American military encourages relevant technological evolution, funds some of it, and smoothes the way for its commercial

introduction. American culture, which tends to be technologically optimistic, underpins these systems, in contrast to, for instance, the European Union, which through formulations such as the Precautionary Principle expresses a deep skepticism of technology. A culture that seeks to match American supremacy in science and technology, therefore, cannot do so simply by increasing research spending or trying to develop a few world-class technical institutes. It must create a network across its culture that understands excellence in this area as an emergent characteristic of excellence across many institutions and domains, a much more difficult task. (Similarly, at a lesser scale sustainable engineering does not require only books and desire; it requires that engineering be valued, that innovation be encouraged, and that appropriate institutions to transmit learning and experience exist.)

But no hegemon, empire, or brand lasts forever. The erosion of power of a great empire is usually not the result of purely external threat; rather, it usually reflects a combination of internal weakness and external pressure. In this light, the Iraq invasion indicates an American lack of sophistication about the real sources of American power and authority. Begin with the observation that America after the Cold War had a dominance that, although displayed in terms of the hard-power supremacy of military and economic might, was primarily cultural. The United States drew brains and capital to itself, and was almost mythically attractive, because of its uniquely open, optimistic, entrepreneurial, and mobile society. The real strength of America was its brand—a strength that, while it rode on, and required, economic and military success, was not limited to these domains. Cultural brands are very different from hard-power technologies: true cultural power is attractive, rather than imposed. The technologies that enable it are very different. In some ways, the management of Coca-Cola's secret recipe, and its subtle variants in different countries, is as critical a technology for American global presence as any military device.

Armies do not congregate and meet each other on the field of battle, as they used to; when that does happen, as in the first Iraq War, American superiority is such that the result is deceptively simple. The geographic simplicity of past wars is far less apparent in current conflicts. Where is the battlefield in Afghanistan, when that conflict is itself part of a contest between Islamic terrorism and the United States that includes attacks on New York City civilian targets and U.S. attacks with Predator unmanned aerial vehicles guided from sites outside Las Vegas, Nevada?

Technological superiority can also be a mixed blessing in an increasingly complex, rapidly changing, cultural and institutional context. Indeed, especially to those who lack experience with the limits of military power, military prowess can be a subtle drug, leading to dysfunctionally oversimplistic policy choices (as discussed in Box 11.1). Just as countries rich in oil or other resources tend to

squander their opportunities, especially in the absence of strong governance, countries rich in military power may be seduced by it into misperceiving their true sources of power (perhaps even more of a temptation for a technologically optimistic culture that will be inherently inclined to seek technological solutions for all challenges). Few successful hegemonic powers achieve their status through military capability alone.

Figure 11.1, provides an initial schematic of the Level I, Level II, and Level III terrain of emerging technologies, military operations, and national security, and illustrates the complexity of the techno-security challenge. Several points are immediately obvious: First, most of the concepts and assumptions that have been stable over much of the past centuries are increasingly contingent. One obvious example is the assumption that the internal activities of states directed against their minorities are not subject to intervention by the international community; another is the assumption that states are the operational entities in international conflicts (Al-Qaeda is an obvious counterexample). This system is highly complex and unpredictable, with different parts interacting in new and uncertain ways; it clearly challenges the sustainable engineer in many different ways. Let's first explore the system at a conceptual level, then turn to several case studies.

Figure 11.1 identifies four major domains. The first, Revolutions in Military Technology (RMT), is the traditional technology dimension, and it is, as the rest of this book has suggested, highly complicated in itself. We will discuss several specific technologies below; the important concept to grasp, however, is that the major challenges arise not merely with specific technologies, which is where we've learned, individually and institutionally, to focus our attention. Rather, it is the cumulative effects and implications of entire suites of technologies, taken across the technological frontier, that pose the analytical and operational challenge. Perceiving and trying to understand the technological frontier as a whole is one of the differences between traditional and sustainable engineering. It is hard enough when one area of technology is evolving rapidly—say, cannon in the late 1400s and early 1500s, or battleships in the late 1800s and early 1900s in Europe. But now with the Five Horsemen it is unclear where military and security surprises might even lie. It is biotech? Cyberwarfare? Something new out of nanotechnology? Although it is difficult enough to understand the implications of a particular technology, the potential for surprise is quite high when the technological frontier is considered, in part because at that level developments are unpredictable. Put another way, the implications of any specific technology can generally be worked in a Level I framework with fair confidence; technological evolution taken as a whole, however, is Level III.

The second realm is that of Revolutions in Nature of Conflict (RNC). Here, a number of institutional, social, and cultural shifts are happening. Thus, for example, one of the long-standing principles of international relations

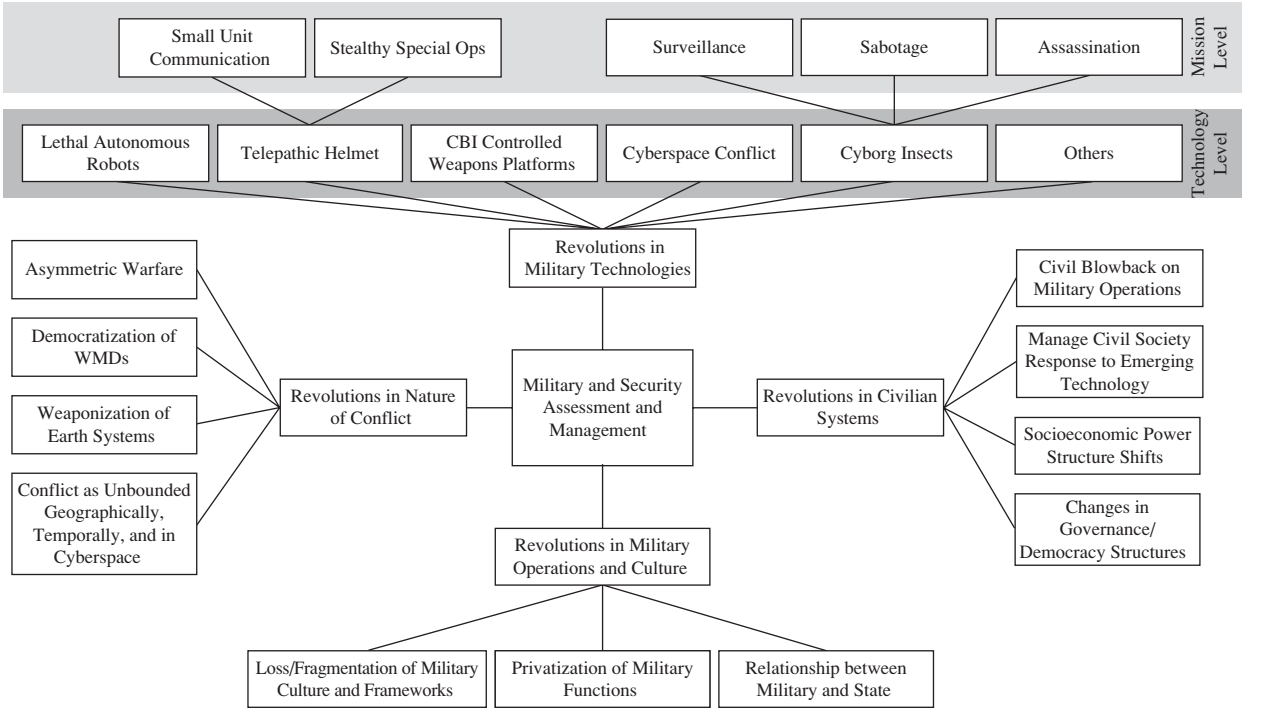


FIGURE 11.1 Emerging Technologies, Military Operations, and National Security. There are four major domains of rapid change in the conceptualization of emerging technologies, military operations, and national security: Revolutions in Military Technologies (RMT), Revolutions in Nature of Conflict (RNC), Revolutions in Civilian Systems (RCS), and Revolutions in Military Operations and Culture (RMOC). Traditional engineering might focus only on the technology level under RMT; the sustainable engineer needs to be familiar with the importance and major implications of all four areas. The drawing may suggest that these are independent, but that is for exposition only; in reality, it is because they are not independent that they are so difficult to figure out. WMD stands for weapons of mass destruction.

since the rise of the nation-state has been the absolute right of the state to do what it wants inside its borders (this was the essence of the Treaties of Westphalia in 1648 that established the international order characterized by nation-states). But now a history of states mistreating minorities—Bosnia, Sudan, Rwanda, the Congo, and many others—has led to strong support for the principle of "responsibility to protect." Under such a principle, countries can intervene in the internal affairs of other countries under certain circumstances. This is a new, poorly defined, and contentious extension of human rights, and is opposed by a number of countries that have difficult internal conflicts (such as China with Tibet, and Russia with Chechnya).

Combat itself has also shape-shifted, from the choreographed battles of the eighteenth and nineteenth centuries in Europe from which much Western theory of warfare derives, to the police actions and insurgencies of Vietnam, Malaysia, Iraq, and Afghanistan, leading to counterinsurgency (COIN)—as the modern model of conflict. Such conflicts are not as limited in time and space as more traditional forms, and they involve new players, such as networked NGOs that cannot be controlled or attacked through traditional means (Al-Qaeda and Aum Shinrikyo, the sect that released poison gas in Japanese subways, are examples). In part, this reflects the success of Western countries, particularly the Americans, in developing virtually overwhelming dominance in traditional methods of war, which makes alternatives, grouped under the label of "asymmetric warfare," essentially the only rational choice for those opposing such states. Levels I and II become entwined here: Weapons developed for combat, such as robots and biological enhancements, can be rapidly adopted by such nontraditional groups and used in conflict against their original developers. This is foreseeable (thus Level II rather than Level III), but the results of such reflexivity are highly unpredictable. (Robot drone attacks on commercial airliners? Hacking of computer–brain interface weapon systems?)

Perhaps the most interesting aspect of this realm is the high level of conflict taking place in cyberspace, often invisible to the public. It is at least possible, and some claim highly probable, that the major players involved—Russia, the United States, Europe, China, India, Brazil, and others—have all compromised each other's Internet systems to the point that any major interpower conflict would immediately result in massive damage to the states involved. (The mechanisms for introducing such chaos would not be the relatively clumsy denial-of-service attacks seen in the Russian assault on Estonia or Georgia, for example, but rather the "back doors" and "logic bombs" that everyone has planted in each other's Internet systems, which would be activated in such a way as to cause optimum damage in case of attack. A delayed "Stuxnet" worm—a software weapon keyed to a specific facility in a specific country—is more likely than generic worms or viruses.) Details, of course, are classified by the countries involved, but it is even possible that a new "balance

of terror" based on "mutually assured destruction" exists in the cyber battlefield. Complicating this is the previously noted trend towards asymmetric warfare: If one wishes to inflict relatively undirected damage on a sophisticated opponent, as many extreme NGOs do, cyberspace is the obvious place to go.

The third realm, Revolutions in Civilian Systems (RCS), is also in a period of rapid and unpredictable change. The interrelationships between RCS and various technology systems nicely illustrate one of the main themes of this book: You don't mess around with powerful Level I technologies without getting institutional, social, and cultural shifts that you didn't expect. There is, for example, the potential for civil blowback arising from technophobia. To take one case, cyborg insects—some sort of robot or cyborg of insect size and function able to conduct surveillance or sabotage—are a powerful counterinsurgency tool that could easily leak back into civil society and be misused, possibly by a rogue military contractor. If that were to happen, society might well react with a ban on such devices, possibly combined with research constraints, with the result that a useful and potentially desirable technology is eliminated from the military repertoire. More generally, suppose that a society is so concerned about the rate and implications of technological evolution driven by military and security needs that it simply seeks to slow everything down—in other words, that it institutionalizes technophobia. Indeed, some have suggested that the European Union and environmental groups are trying to do this with the Precautionary Principle. Over time, such a society will become less competitive with those that continue technological development. In other words, a military technology policy that may be extremely effective in a Level I sense across the board could well jeopardize the long-term viability of the security of its culture if it leads people to turn against technology as a whole.

Such possibilities do not exhaust the categories of potential impact, which would in addition include Revolutions in Military Operations and Culture (RMOC). This is not just of interest to the military or to the security apparatus. A big reason some countries are coup prone and others are not is culture. In this case, emerging technologies, combined with some of the changes mentioned previously, raise a number of challenging scenarios. For instance, the military culture in most developed countries assumes a professional, highly trained core around which volunteers or draftees are assembled. But operations in Bosnia and Afghanistan are more like policing operations, where the object is to protect civilians and the built environment, rather than to blow it up looking for bad guys. Policing and combat require very different kinds of technology, training, and institutional cultures—one reason why good soldiers are seldom good policemen, and certainly not both at the same time. Technology that is very useful in combat, such as machine guns or high explosives, may be very counterproductive if used in policing operations, especially in civilian crowd control.

But it gets more complicated. Consider, for example, the growth of robots in Iraq, where according to Peter Singer in his book *Wired for War*, the United States, which had 0 ground robots at invasion time and only 150 by the end 2004, grew that number to 2,400 by the end of 2005; 5,000 by the end of 2006; and had 12,000 by the end of 2008. Even more explosive growth has occurred in the number of aerial robots, or UAVs. One result is a profound culture conflict, as officers who have absorbed military culture for many years as they worked their way through the ranks find themselves competing with gamers, and military personnel who are in active combat situations in the Hindu Kush are augmented with remote control operators who go home to their houses in American suburbs when their day is done. (Is one effect of UAV technology operated in such a manner to convert American suburbs to legitimate military targets?)

On top of that, many functions previously carried out by active duty military personnel, including combat roles such as guarding convoys and running military prisons in non-U.S. locations, have now been privatized, turned over to companies such as Halliburton and Blackwater (renamed Xe after some unfortunate incidents that involved the shooting of civilians). This has huge implications from the perspective of emerging technologies. It is one thing to have cyborg insect technology under state control, but quite another should it dribble back to civil society through quasi-governmental private firms, with neither the culture nor some of the legal protections that might otherwise apply. This is especially the case because private military contractors do not necessarily owe allegiance to a particular state or culture, and can if they choose shift themselves to other jurisdictions.

This mapping, really just a conceptual schematic at this point, illustrates some of the complexity that a serious consideration of emerging technologies necessarily involves. The combination of RMT, RNC, RCS, and RMOC is not something that is avoidable if one seriously contemplates issues of technology, national security, and sustainable engineering. Each realm is changing, each is contingent, each is unstable, and they are all coupled to each other. This is an extremely volatile situation to introduce new technologies into, and one may well ask why one would do so, given that many technologies of sufficient power to be of interest militarily have at least the potential to be deeply destabilizing to existing economic, social, and technological systems. Examples might include the possibility that military radio frequency identification (RFID) sensor systems, insect robots, and small-scale cyborgs are shifted from theater intelligence to domestic intelligence; that telepathic helmet technology transitions from a small unit communication enhancement to a nonintrusive thought detection device in civil society; or that warrior enhancement technology results in radical life extension for selected civilian populations.

It is therefore worth reviewing some of the reasons why such technologies are developed and deployed before diving into a few scenarios. To begin with, effective military technologies enable more potent projection of national power, and in some cases (e.g., ground-based bomb squad robots) directly save combatant lives. They work and they save lives; this makes them extremely hard to oppose, at least in terms of development and deployment in combat situations. Moreover, many emerging military technologies substitute capital for labor, meaning that one gets more performance with fewer individuals. This is valuable because, just as in the civilian world, demographic shifts to older populations in many countries will make maintenance of military organizations, heavily dependent on younger individuals, more difficult absent technological fixes.

Once they are out of the box, however, technologies that are of use in the wider civilian world tend to spread, especially where they don't require substantial expertise or infrastructure to replicate. Moreover, given the nature of military competition, many important technology systems are developed in classified environments: there is no opportunity for a wider public dialog before they are introduced. In addition, global technological competency raises further issues: That an individual country or society may eschew a technology does not mean that its potential competitors will. The Soviet Union, for example, continued to develop biological weapons through its Biopreparat program long after signing agreements not to, and research on poison gas formulations is necessary to develop effective countermeasures. The difference between defensive research on a particular technology and the development of offensive capabilities is often very difficult to determine, especially in classified programs.

This poses a number of issues for sustainable engineering. It is useful to remember that, while one might wish away the security and military applications of technology, in the real world, with real problems and conflicts, this is simply not possible for the engineer. For this reason, the superficial answer—"just say no" to difficult technologies and choices—often simply fails. The sustainable engineer in this domain must instead try to understand this terrain as best one can in real time and manage it as well as possible given existing and foreseeable constraints. A few examples combined with the discussions of technology in the previous chapters may be of some use in this regard and provide a few instances of some of the technologies that might have come from the military R&D machine. Remember, however, that these technologies are still under development. Thus, although they may be funded and some of the preliminary science looks promising, they still should be regarded as scenarios rather than done deals. On the other hand, these descriptions are necessarily based on public sources, so they don't include what might already have been accomplished on the classified side.

Case 1: Cyborg Insects and Other Surveillance Devices

This set of technologies involves creating either robots with the size and functionality of insects, or implanting electronics in biological insects, to create cyborg insects. Research that has already been presented in the literature (as of 2009) includes successfully implanting electrodes and a radio receiver on a giant flower beetle, allowing that insect to be controlled during flight. Moth pupae have also had silicon chips implanted in them; the moths then retain the chip and connections as they pupate. Research on technologies to network cyborg insects into grids, thus dramatically increasing their computational and surveillance capability, is also underway. (This is an extension of existing work on interconnected systems of coffee can size sensors; in a system called "Wolfpack, "for example, each sensor alone was dumb, but the interconnected Wolfpack was pretty smart.) Although much attention has been focused on flying platforms such as large beetles, the cyborg insect zoo under consideration includes variants of spiders, high-jumping insects such as grasshoppers, moths, and bees with detachable payloads, and the like. Serious technical challenges include designing energy sources for such small platforms, especially as guidance, surveillance, or attack modules become part of the payload, and miniaturizing functional equipment so that it can be fitted on an insect-sized carrier.

This set of technologies is interesting in part because the Level I and Level III implications potentially tend strongly in opposite directions (Figure 11.2). From a Level I perspective, the alignment among technology, goal and policy is pretty clear. A surveillance cyborg insect system would be a very powerful military tool in two environments that form the bulk of current direct combat commitments for the United States: urban low-level conflict and COIN operations (the focus is on U.S. military conditions because much of this technology is being developed in the United States). Pressures to develop this technology are therefore quite strong, because the Level I implications are so clearly positive. Moreover, cyborg insect technology is obviously a generic platform. While most discussions involve surveillance functions, there is no reason cyborg insects can't be weaponized with either nonlethal or lethal capability. In either case, it is arguable that the technology is potentially more ethical than existing alternatives. UAV strikes from Predators, for example, usually involve a small but still undesirable amount of collateral damage ("collateral damage" is the euphemism for destruction of property and the killing or injuring of civilians). A cyborg insect system could identify a target, and immobilize or otherwise attack just the target, reducing civilian and property collateral damage considerably. Even in a purely surveillance mode, use of this technology

Policy Response / Technology Level	Goals and Effects	Policy Response
Level I: Military effectiveness	Reduce collateral damage and increase operational efficiency in COIN and urban operations.	Goals and technology strongly align; therefore, adopt technology.
Level II: Adoption in civil society	Implement in domestic society to protect civilian populations from terrorists and, through mission creep, criminals.	Goals and technology strongly align; therefore adopt technology.
Level III: Social, cultural, and earth systems effects	Technology likely to be privatized, with substantial reduction of privacy and contribution to "soft" or "hard" security state (vs. free speech/individual freedom democracy).	Optimistic goals likely to be undercut as those in power adopt cyborg insect technology to their own ends; Level I and Level III implications potentially in fundamental conflict.

FIGURE 11.2 Policy Response Matrix: Cyborg Insects. The technology analysis matrix for cyborg insect technology shows an example of a situation, not uncommon with military technologies, where Level I positive benefits may distract one from more problematic Level II and Level III situations (note that the Level I benefits will be fairly clear, whereas the Level II and especially Level III implications will be far more hypothetical, given the much greater complexity of higher level systems).

could produce much more detailed information on targets, which in turn can also significantly reduce collateral damage.

But Level II and III are different matters. Level II might involve the use of cyborg insect technology in civil society, a clear and likely extension of the technology given the radical privatization of military operations that has also characterized recent conflicts. The implications for personal privacy, for example, could be both widespread and devastating. Over time, the balance between public safety and security, and privacy and independence, that most societies vacillate around could be significantly pushed towards the "intrusive state" end of the scale by such technologies. Grid technology would only exacerbate the reach of an intrusive elite or state. Given how relatively obvious this potential is, and how well it fits with current surveillance strategies such as the ubiquitous cameras in the United Kingdom, this

is a foreseeable, but complicated, Level II scenario. Note here how the different aspects of security identified in Figure 11.1 interact. If warfare were not being privatized so significantly, perhaps the introduction of such technologies into civil society would be easier to control; if fear of terrorism and crime weren't already pushing traditionally democratic states such as Britain and the United States more towards surveillance and privacy intervention, one might not be quite as concerned about potential domestic abuse of such technologies.

But if this technology spreads, the full implications are hard to determine. In the never-ending conflict between totalitarian and open government, technologies play unpredictable roles. Twitter, for example, has served as a technology that spreads information about domestic unrest around the world, and thus seems to push towards openness; cyborg insect technology, on the other hand, would likely push strongly otherwise. Moreover, cyborg insects are a technology platform, not just a surveillance technology: one could as easily fly a swarm of cyborg insects into a nuclear plant with the idea of destroying it as one could collect data on a rival or opponent. Imagine how ugly politics could get when political parties have their own fleet of cyborg insects. Conversely, one could imagine gridded cyborg insect populations that could provide real-time disease or environmental condition sensing systems, potentially providing great benefits to society.

What is clear from this example is that some technologies have Level I military operation implications that are very different from their potential Level III cultural and earth systems ones. From a military operations perspective, this is a very interesting and potentially useful technology. But once it spreads to civil society, it could become an important force tending against an open, democratic society, with the end result that security is saved at the cost of what made the society worthwhile—a Pyrrhic victory indeed. In cases such as this, therefore, whether a technology should be developed and deployed involves a difficult balancing act across categories by the sustainable engineer. What the right answer is cannot be determined without much more information and study, and will become apparent only in real time because of the unpredictability of the system. Which hypothetical might actually develop can be known only as a result of a yet unplayed out interaction among technology development—economic feasibility; law and regulation; social and cultural norms (how much surveillance do most people want?); and a slew of other variables. But at the least the analysis warns one to be very careful of confusing categories and implications, and to be very aware of trying not to let the obvious advantages at Level I blind one to the need to identify and manage serious potential disadvantages at Levels II and III.

Case 2: Telepathic Helmets

This is also a composite technology. It integrates strands from a number of different areas of scientific research, most obviously drawing on work that demonstrates the ability to determine what noun, image, or moving image is being thought of by an individual via noninvasive interpretation of brain activity and emissions, combined with greatly expanded computer-based modeling of brain architectures. More fundamentally, this reflects the accelerating progress in both computational power and neuroscience that has been discussed elsewhere; in 2009, for example, IBM announced a new supercomputer simulation where the number of (algorithmic) neurons and synapses exceeded those in a cat's brain. The benefits of a computer–brain interface (CBI) helmet, which would read a soldier's brain waves and transmit that content to other members of a small unit, thus enabling such units to act more as a cohesive whole and share critical information instantaneously and easily, has led the U.S. Army to fund a multimillion dollar exploratory contract to researchers from Carnegie Mellon University, the University of California at Irvine, and the University of Maryland. Think of it, in short, as a telepathic helmet—albeit a primitive one for the foreseeable future. In practice, it is highly likely that the telepathic helmet would also integrate information from "smart" technology systems scanning the battlefield and prioritizing threats and target opportunities—an augmented cognition (augcog) function—but for simplicity's sake we will only discuss the telepathic technology.

Now, it's true that toys controlled by thought are already available; one, for example, allows control of a ball in air by means of thoughts affecting the air stream holding it up, and headbands allowing gamers to mentally interact with their avatars can be purchased for less than $300 as of 2010. But it is a long road—those working on the technology suggest decades—from such devices, which use relatively high-level and coarse brain function patterns, to being able to do the fine work required for a telepathic helmet. So for now it's very appropriate to view telepathic helmet technology as a scenario. Nonetheless, it is a very interesting one, as the policy response matrix analysis of this technology in Figure 11.3 suggests.

As in the previous case, the Level I analysis is relatively easy, both to understand and to evaluate ethically. The technology enhances performance of small combat units, especially where face-to-face and voice contact is difficult (as during special operations, for example). It also may enhance unit cohesiveness. It will thus increase performance, reduce unnecessary mortality, and quite possibly reduce collateral damage, especially if combined with other technologies (e.g., telepathic helmet linkup to cyborg insects and augcog applications). So at a Level I, it's a go.

Policy Response / Technology Level	Goals and Effects	Policy Response
Level I: Military effectiveness	Enhance performance of small units in combat.	Goals and technology align; therefore, adopt technology.
Level II: Adoption in civil society	Protect civilian populations from terrorists and, through mission creep, criminals by enabling noninvasive distance capture of thoughts.	Implement technology, but technology alone may not lead to achievement of stated goal.
Level III: Social, cultural, and earth systems effects	Elimination of privacy and social instability as subconscious and unfiltered thoughts become explicit. Fundamental changes in language and communication. Enable security state versus democracy.	Level I and Level III implications potentially in fundamental conflict. Effective military technology dramatically alters the culture and society that military action was undertaken to protect in the first place.

FIGURE 11.3 Policy Response Matrix: Telepathic Helmet. The technology analysis matrix for telepathic helmet technology displays a similar pattern as with cyborg insects—clear Level I benefits with more problematic but more hypothetical issues at Level II and Level III. Because technologies for military or security applications are generally not developed unless they have at least potentially significant advantages for military operations, such patterns are characteristic of military technologies. Indeed, they are characteristic for technologies as a whole.

But note the implications of the technology. If an engineer can build a helmet that knows what you're thinking, she can also build devices that can read your thoughts from a distance; if she can do that, the user of the technology can know your thoughts without you knowing that you're even being read. (This possibility assumes somewhat heroic advances in our understanding of cognitive function, as well as computational power, but given accelerating advances in those fields, it is not implausible.) From a Level II perspective, this might appear to be desirable: Airport security, for example, would gain considerable punch with such technologies. But, as with other Level II technologies, no technology by itself is going to accomplish the systems goals characteristic of Level II. It is possible that such "brainspy" technology could indeed enhance security. However, it is also possible that, if we become too reliant on technology to the detriment of other, less technocratic practices—a well-trained security observer,

for example, and human infiltration of terrorist infrastructures—we would in fact reduce security overall. Moreover, it is not at all clear that just because a person thinks something violent, that person is in fact violent. Research on the brain indicates that we may think passingly of many things that are overridden by other, balancing networks within the brain. To think, "I want to hit you for what you just said," is not the same as actually doing it. To the benefit of the community and fellow humans, it appears that we think a lot of things we never do.

This leads to some Level III speculation (anything said about Level III effects so far in front of the technology itself must be speculative). Consider what a telepathic helmet will communicate: not just nouns and verbs, but images, moving pictures, fragments of sound, perhaps memories of smells and touch, and emotional packages. It is an extraordinary broadband communication system that the engineer unlocks here, one that humans have never, in their entire history, had to contend with. What happens to language? What have the technologies of email and Twitter done to the English language already—to spelling, to punctuation? What have Facebook and other social networks already created in terms of composite communication platforms? Note that even at this relatively primitive beginning of social network technology it is increasingly a world of multimedia transmission, with words and punctuation significantly modified by the requirements and capabilities of that technology. What would a full-bore telepathic connection transmit? Language would surely and unpredictably change.

And what about the self? Although the boundaries are different in every culture, we are all used to three bins: a "public" self that is what the world sees; a "private" self that is limited to us and our friends; and a "no one else is ever, ever going to know what I'm thinking" self that is that part of our mind that is never shared and that with current technology is in fact impossible to share. This changes with the telepathic helmet. All of that is now potentially sharable. The bandwidth is open. College romance will never be the same. Nor will employment interviews, or arguments with your parents or friends. And, of course, if the technology develops as it might, everyone else has access. Moreover, being able to understand the cognitive system well enough to capture and transmit such information opens up other possibilities: the potential for hacking in such systems, for eavesdropping, even for rewiring other's brains through reverse transmission. After all, the printing press (technological change) and rapid increases in literacy (social change) led to a much more educated, much more informed society, where the written word was privileged in many ways. It is somewhat disconcerting to see the next phase at work and not be able to know what the effects might be.

There is no way of knowing which of these possibilities may start becoming real and which will, with experience, turn out to be laughable paranoia. In each instance, it is possible to suggest positive aspects of these scenarios. Many

people, for example, might benefit from a technology that makes their loneliness less acute. Actually, it is highly likely that whatever does come of this technology will be something that, because of our preconceptions and inability to sufficiently appreciate the contingency of our assumptions, will be totally unanticipated and, yes, unpredictable. What is clear, however, is that the Level III implications of this sort of technology are extraordinary, complex, and beyond most people's thinking at this point. Given that telepathic helmet technology, like any technology, carries all these implications simultaneously, this is dangerously shortsighted. It is not that we shouldn't appreciate Level I, but that we should also learn to be looking for, and trying to understand, Levels II and III simultaneously. Just because something sounds like science fiction, doesn't mean it really is fiction. It may be deploying in Afghanistan tomorrow.

A final important observation for sustainable engineering. Note that the Level I effects of these technologies are all obvious and very positive. That the goal, the technology, and the policy are coupled and self-enabling makes Level I understandable to all. The Level II and III effects, however, are mostly hidden, and reflect networks of factors that most people don't understand, much less know how to manage. As a result, there almost always will be strong pressure to implement technologies, military or otherwise, where the Level I implications are clear and positive, regardless of Level II and III implications. A good sustainable engineer will take this into account, not by adjuring technology, but by supporting the development of institutional capability, in governments, professional organizations, and NGOs, to explore, understand, and react appropriately to Level II and III changes as they can be identified.

Case 3: Lethal Autonomous Robots

Case 3 also involves a suite of technologies, in this case with two relevant characteristics: "Lethal" means that the robot is equipped to kill human targets, and "autonomous" means that the decision to kill is the robot's, not a human's (that no human intermediary is in the loop between target identification, verification, and elimination).

"Robot" sounds like the most obvious of the lot, but it is not. Most people agree that a machine such as the Predator or Raven UAVs, the Talon, PackBot, Warrior, or Swords tracked platforms used in Iraq and Afghanistan, or the unmanned underwater vehicles currently deployed by the U.S. Navy are robots. But some people consider some land mines to be robots: They are inert until they sense the proper triggering conditions, then they jump, or explode, or do whatever else they are built to do. And what about a grid of surveillance/attack cyborg insects? Each cyborg insect taken alone may be too dumb to be considered a robot, but the grid as a whole may be quite intelligent. And what should we call a weapons platform that is wirelessly connected

directly into a remote human brain? (In recent experiments, a monkey at Duke University with a chip implanted in its brain that wirelessly connected it to a robot in Japan kept the Japanese robot running by thought, so that the robot was in essence an extension of its own physicality.) Even now, the Aegis computer fire control system deployed on Navy vessels comes with four settings: "semiautomatic," where humans retain control over the firing decision; "automatic special," where humans set the priorities but Aegis determines how to carry out the priorities; "automatic," where humans are kept in the loop but the system works without their input; and "casualty," where the system does what it thinks necessary to save the ship. Similarly, the Counter Rocket Artillery Mortar system (CRAM) is a computerized machine gun system currently deployed to defend against rockets and missiles that humans cannot react to in time.[3] Is "robot" a type of artifact, a type of capability, or a certain level of computational competence? Or does this discussion make the interesting point that emerging technology also makes our language itself contingent, and if we're not careful, we'll end up trying to use obsolete concepts and words to describe a world that has already moved on?

To return to the technology itself, lethal autonomous robots (LARs) are controversial at every level, but the discussion to date is often characterized by category confusion. Let's try to sketch out some of this difficult ground (Figure 11.4).

To begin with, why should such a technology in any form be deployed? Here, we have serious coupling between Level I and Level II and III issues. The immediate response is, "To save soldiers' lives." And this is a true Level I benefit. Robots, whether lethal and autonomous, or simply robotic, do save the lives of soldiers. Many explosive devices in Iraq and Afghanistan that might otherwise have killed and maimed soldiers have been identified and eliminated by robots. But this is in some ways only begging the question—after all, in World War I generals thought little of killing 100,000 men at a go by sending them into the teeth of concentrated machine gun fire. So we need to go a little deeper.

The U.S. military in particular faces a stark dilemma. It is charged by its society with being able to project force anywhere around the world, under virtually any conditions. On the other hand, American civilians for a number of reasons, such as much smaller family size (so that the loss of a child, never easy, is even more traumatic), have become averse to any casualties. The military is thus supposed to project power without American soldiers dying. Moreover, the long-term demographics are not good: Americans, like other developed countries, are looking at an aging demographic, as discussed in Chapter 2. There are less young people to fill boots on the ground (this is why

[3]These examples come from the book *Wired for War* by Peter Singer.

Policy Response / Technology Level	Goals and Effects	Policy Response
Level I: Military effectiveness	Project force without endangering humans; substitute machines for increasingly scarce people.	Goals and technology align; therefore, adopt technology.
Level II: Adoption in civil society	Create technologies that are able to substitute for humans in undesirable and low-paying jobs, especially in aging populations (e.g., Japan).	Implement technology, but technology alone may not lead to achievement of stated goal.
Level III: Social, cultural, and earth systems effects	Substitution of robots for humans, and integration of robot with human, creates human varietals with greatly expanded capabilities (e.g., operate in space environments). Major but unclear implications for human and social evolution.	Level I and Level III implications potentially in fundamental conflict. Effective military technology, but effects of integrating the human with the robotic unpredictable at Level III.

FIGURE 11.4 Policy Response Matrix: Autonomous Robots. This is a more complicated case than the previous two, in that the Level I implications are not necessary as clearly beneficial as in previous cases. (For example, what if the robots aren't able to adequately differentiate between enemy combatants and civilians?)

the U.S. Defense Advanced Research Projects Agency, or DARPA, funds life extension technologies—if there are less young people, you've got to keep the ones you've got younger longer). The net result is you've got to design towards a state where you get better military productivity, with productivity measured as mission accomplishment per soldier lost.

Robots are a major technology to achieve this goal. It's not just about saving soldier's lives, a Level I technology. It's also about building the capability to continue to project power with less casualties, a Level II goal, and to do so because culture and society are changing to make fatalities, whether soldier or civilian, less acceptable, a Level III trend. Moreover, there is a more subtle Level III effect occurring here: The pressures that current risk aversion and demographic trends create on the military are exactly the pressures that the early Industrial Revolution created—and the technological response is the same. Just

as the history of capitalism from the beginnings of the Industrial Revolution has been characterized by increasing efficiency and production arising from the replacement of human effort by machines, so too is LARs technology a long-term substitution of capital for labor to enhance military efficiency and a concomitant increase in the productivity of the individual. Note how this latter trend pushes one towards *autonomous* military robots: You reduce death if you replace a soldier with a robot, but you don't augment individual productivity until you deploy autonomous robots. With existing UAVs such as the Predator, for example, you still need one human per machine (actually, each machine requires a team to keep it operational), but you can have one human supervising many machines if you deploy LARs. And unlike the previous technologies, LARs are not just a scenario; they can be deployed quite quickly (and by some definitions, systems such as Aegis are LARs already).

But there are concerns about deploying LARs, even at Level I. For one thing, the laws of war[4] require that combatants adhere to two requirements: discrimination and proportionality. A combatant must be able to discriminate between enemy combatants and civilians, and is not allowed to target civilians. A combatant is also required to use only such force as is necessary for self-protection and to accomplish the mission (the proportionality principle). Opponents of LARs argue that robots cannot possibly discriminate between civilians and combatants, or exercise discrimination, and are therefore not permissible under the existing laws of war. Proponents of LARs, on the other hand, argue that this position is, itself, a product of discrimination, in the form of a belief that the way soldiers behave now is in fact better than robots can be programmed to behave. Robots in this view need not be perfect, they just need to be better than humans are under similar circumstances—and proponents note that humans do not function nearly as well under the stress of physical combat as the opponents of LARs seem to assume. Indeed, the assumption of those opposing LARs that human soldiers behave better under actual combat situations than robots might is seldom defended and, given the realities of combat is at least questionable. For example, it is not trivial for a warfighter to differentiate between combatants in a counterinsurgency environment—the same person may be at one time a farmer and at another time an enemy combatant. So in this case, the Level I arguments are somewhat mixed.

[4]A deep discussion of the laws of war is beyond the scope of this chapter. In general, however, it should be noted that there are three main branches of the laws of war: (1) legal justification for engaging in war in the first place (*jus ad bellum*), (2) legal conduct during war itself (*jus in bello*), and (3) international agreements on military conduct and technology (such as the Geneva Conventions and other treaties). In addition, domestic law may govern aspects of declaring and conducting war. Moreover, in part because of the unprecedented changes in RMT, RNC, RCS, and RMOC discussed above, many are questioning whether the laws of war as currently constituted, which primarily reflect Western history and concepts of war, are still viable in whole or in part.

The Level II and Level III arguments are also complex. For example, a robotic force, including LARs, would clearly save lives, but the potential for loss of life is perhaps one of the most powerful reasons why societies, especially democracies, tend to avoid war except as a last resort. Minimizing the loss of life, therefore, might increase the possibility that war would be chosen as a viable response to a situation that could otherwise be resolved by negotiation. LARs also increase the distance between those who decide to use force against other human beings and the physical results of that decision. Even today, flying a Predator and watching the flash that destroys a target is a very different psychological experience than shooting an enemy at close range in Afghanistan. At the ultimate, killing an enemy becomes a video game (as in the science fiction classic *Ender's Game*); indeed, on the Internet today one can find a genre of films called "war porn," where images of real death and injury are juxtaposed with humorous dialogs or popular music.

But these technologies also lead in another Level III direction, which is the integration of the human and the robotic in such a way as to create families of human varietals designed for different environments and behaviors. After all, Adam Smith's division of labor was a major factor in the Industrial Revolution and the rise of the modern world; today, the homogeneity of the human species can be seen as an enormous inefficiency given the challenges and opportunities of the Anthropocene. Humans as wetware, for example, are terribly designed for space environments, or for Mars, or for highly complex cognitive tasks; but human/robot combinations could be much more effective in space travel. Humans as wetware have certain cycle times and perceptual bounds; coupled with robots either directly (via wireless implants, for example) or indirectly via plain old video controllers (many of which have been ordered for Iraq and Afghanistan because they enable soldiers familiar with such devices to more easily control robots), human cognitive systems can expand their boundaries dramatically. And military development of LARs and other robotic options would make such scenarios much more probable, because the technology that enables a soldier in relative safety to couple to a robot in a combat zone is not that different from the technology that enables the earthbound scientist to couple to robots on the moon.

11.3 CONCLUDING THOUGHTS

There are several important sustainable engineering principles illustrated by this chapter. Most obviously, perhaps, the Level I, II, and III categories do seem to offer a useful framework for thinking about technologies, their implications, and potential policies. In addition, this discussion has illustrated the importance of the domain of emerging technologies, military operations, and national security for sustainable engineering. It is not only critical to under-

standing the world as it is now and as it will be, but it is where much of the future is being forged. It has also illustrated one of the complicating factors about emerging technologies: sustainable engineering must obviously include a feel for the issues raised by such technological evolution, but, at the same time, the lines between science fiction and science fact are a lot blurrier than one might wish. It's not that sustainable engineering can develop the ability to predict how complex adaptive technology systems will evolve over time—they are inherently unpredictable. But clearly it is possible to think more constructively about potential futures and to use scenarios to explore what they might mean, and develop option spaces that enable society to adjust as implications of emerging technologies become apparent. We need to know when things can be safely put off until tomorrow and when something is happening that we ought to be concerned about now. We also need to be able to shift our institutions and their cultures more effectively and agilely. That capability is going to require some work, but learning to think about these systems more rigorously and with deeper understanding is surely a prerequisite. Once again, we find a need for sustainable engineering not just as a professional capability, but as a critical social competency—and for the sustainable engineer not to be just technically competent, but to aspire to leadership in his or her society.

Exercises

11.1 You work for a government military defense laboratory, and one of your scientists has just developed the computer, data, and hardware systems necessary to directly couple the brain of a soldier to a major weapon system. Create a technology analysis matrix for this technology.

11.2 As the first scientist leaves your office, a second enters to announce that a pill that eliminates fear in humans has been developed and is ready for deployment. As you think about this technology, you begin to worry that even the Level I implications are complicated and involve both benefits and drawbacks. Write a memorandum for your boss, the head of your laboratory, outlining the pluses and minuses of this technology in combat (that is, ignore Level II and Level III scenarios).

11.3 Nice try. Your boss calls you up and wants to know what the implications of such a pharmaceutical might be if it were, in her words, to "get loose" in civil society. Write a second memorandum back to her discussing Level II and Level III implications of the technology.

11.4 Apart from the United States, the BRIC countries—Brazil, Russia, India and China—have started to emerge as the most powerful countries in the world. Given that the 'hard power' of BRIC is low compared to the United States, which spends nearly as much as the rest of the world combined, how would you characterize the forms of soft

power demonstrated by each of these countries?

11.5 The U.S. Army has proposed a technology package that would bioengineer soldiers to be permanently altered to have 200 percent greater strength, nervous system function, cognitive capability, and skeletal strength. You are the senior Pentagon analyst for emerging technologies. Write a one-page memo to the Secretary of Defense outlining the Level I, Level II, and Level III potential implications of this set of technologies. You can use bullet points if you so choose.

Annotated Bibliography

1. There are several general military histories that the interested student will find useful. If you only have time for one, perhaps the best overall general text is John Keegan's 1994 book, *A History of Warfare* (New York: Vintage Books). Other excellent general histories include Geoffrey Parker's 2005 edited volume, *The Cambridge History of Warfare* (Cambridge: Cambridge University Press), and Max Boot's 2006 history, *War Made New* (New York: Gotham Books). Note, however, that these are all histories that tend to focus primarily on the Western history and practice of warfare, in part because for several centuries that has been the dominant form of conflict globally. Wars that do much damage, but occur in other contexts, such as many of the African conflicts of the twentieth century, are not covered in as much detail.
2. The interrelationship of modern technology and modern war, although still evolving, has been the subject of an increasing number of books. Among the more accessible ones is Peter Singer's 2009 text, *Wired for War: The Robotics Revolution and Conflict in the 21st Century* (New York: The Penguin Press); another excellent text on the subject of the technology and revolutions in military operations and institutions is MacGregor Knox and Williamson Murray's 2001 edited volume on *The Dynamics of Military Revolution, 1300–2050* (Cambridge: Cambridge University Press). William McNeill's 1982 book, *The Pursuit of Power,* is an excellent study of the development of military institutions and technology, primarily in Europe; it is noteworthy because it discusses the effects of long-term population dynamics on willingness to wage war and the coevolution of complex, expensive military technologies and institutions, and industrial capitalism, in ways that other texts don't.
3. It often is quite enlightening to go directly to the source, or in this case, the field manuals. The most interesting ones right now are perhaps those recently issued by the U.S. military reflecting experience in Iraq and Afghanistan: *The U.S. Army Stability Operations Field Manual, No. 3-07* (a 2008 version was published by the University of Michigan Press), and *The U.S. Army Counterinsurgency Field Manual, No. 3-24* (published in 2007 by the University of Michigan Press).

CHAPTER

12 The Macroethics of Sustainable Engineering

Fantasy, abandoned by reason, produces impossible monsters; united with it, she is the mother of the arts and the origin of marvels.

—Francisco de Goya, Spanish artist, 1746–1828

12.1 INTRODUCTION

Just as surely as artists bring their visions from infinite possibility and create marvels, so do engineers, with the critical difference that in an engineered world people's lives and property often depend on how well the engineer has done her or his craft. In some sense, then, ethics and engineering have always been engaged with each other. Engineers, after all, differ from scientists in that they create designs, products, infrastructures, and these days virtual realities and biological systems that never before existed and that without the engineer would not exist. That is, engineers make choices about what will be, and what will not be, and they make choices about what forms, styles, and cultural leanings are expressed in our environment. When bridges on the Long Island Expressway were designed to be low so that buses carrying poor people were not able to pass under them and thus get to Long Island

beaches from New York City, while the cars of wealthy individuals were easily able to, it was not just a design statement—it was an ethical statement. When Baron Haussmann designed a new Paris in the 1860s for Napoleon III and created the new, spacious boulevards that for moderns are the very essence of that beautiful city, he did not do so just for aesthetics or for grandeur, although both resulted from it. Nor did he do it just to destroy the past, although it is estimated that some 60 percent of Parisian buildings were affected in some way by his sweeping reengineering of the city (indeed, there was considerable angst at the time because of the beauty being swept away). Napoleon III hired him, and his design was intended, to enable troops to rapidly disperse throughout the city in case there was another revolutionary uprising. (In 1848, there was a wave of revolutions throughout Europe; in Paris, a major tactical advantage of the revolutionaries was the old neighborhoods with narrow and difficult to navigate streets, which were easily blockaded by street fighters.) The Paris you love today was built to ensure that workers could be crushed by the State.

Creation is often an ethical act, and engineering is no different. This is especially true of sustainable engineering, where the engineer is being asked by society to give voice to vague and mythic yearnings; and it is especially true in an age when earth systems, including the human itself, are becoming design spaces. Sustainable development as set forth in the authoritative World Commission on Sustainable Development Report, *Our Common Future* (Oxford: Oxford University Press, 1987), establishes two basic ethical principles: Environmental quality is important, and an egalitarian social structure, where wealth and assets are shared relatively equally in and among generations, is preferable to other alternatives. These are not objective observations—statements about what "is"—but normative commands—statements about what "ought to be." In all this, how can one think constructively of ethics?

Let's begin simply, by defining ethics. From an academic perspective, ethics, which is sometimes called "moral philosophy," is the field of inquiry that tries to understand questions about what is good (and bad), what is right (and wrong), and concepts such as justice, virtue, and goodness. It has a number of subfields, such as "applied ethics," which studies the use of ethics in particular situations and professions, and "descriptive ethics," which looks at how people actually behave. More recently, there is a growing scientific interest in ethics, which asks questions such as: "How do certain kinds of ethics help individuals or groups survive?" "Are there neurological bases for ethical behavior?" For most of us most of the time, ethics is simply the effort to assure that we're doing the right thing, and making the appropriate decision under the circumstances.

Understanding the relationship between law and ethics is important. The two are related, but they are not the same. For example, much activity that might be considered unethical is not illegal. Harsh punishment of a child, for example, is often not illegal unless it is excessive, because as a general matter

the law assumes parents have broad rights when it comes to children. That doesn't make it ethical. Falsifying data in an article would be considered by any engineer to be unethical, yet it is not, absent some further harm, illegal (falsify data in a grant application to a government, however, and you may well run afoul of statutory provisions). Indeed, sometimes the law and ethics conflict: Changes in a design being built in a wetlands in many jurisdictions, even if the purpose and effect is to protect the local ecology, would be illegal if they contravened the environmental permit issued for that activity. It is generally a good idea, therefore, to be explicitly concerned with *both* the legal and ethical considerations applicable to one's engineering activities, and independently assure compliance with both: The assumption that because one is behaving ethically one is also behaving legally is not necessarily true.

It is also important to differentiate between ethical standards and the ways those standards may be expressed in different cultures. For example, consider two burial practices. In one case, the brain and parts of the body of the deceased are consumed by relatives; in another case, the body is treated with chemicals, placed in a large box, and buried. Europeans labeled the first practice as cannibalism, and held it to be not just unethical but evil; they, of course practiced the second. On the other hand, the Papua New Guinea tribes that practiced ritual cannibalism at the death of a loved one would have found the Europeans' treatment of their dead equally inexplicable and sinful. In both cases, however, the culture was expressing the deep feeling of loss and mourning caused by death of a loved one, and was honoring the dead—in other words, respecting the ethical standard of mourning a death as a rite of human existence. But they had very different ways of expressing that standard, and those differences led others to condemn them as unethical when what in fact was being expressed in both cases was the same basic ethical principle.

12.2 FRAMING ETHICS

One sometimes hears professional ethicists making assertions such as "engineers have no ethics," or "engineers need to be trained in ethics."[1] Such statements are simply wrong, and usually reflect the fact that engineering education tends not to have specific classes in ethical theory and classical moral philosophy. But this does not mean that engineers are not highly ethical. Conversely, that there is a robust theory and practice of ethics already operating in engineering as a profession does not mean that both existing theory and practice do not need some difficult rethinking given the anthropogenic world and the Five Horsemen that we have discussed in earlier chapters.

[1]The author is the Lincoln Professor of Engineering and Ethics, and as such has heard a number of statements to this effect from professional ethicists, primarily because they seem to confuse a lack of formal training in ethics as an explicit subject of study with an underlying lack of ethics.

- Personal
  - Everyone has ethics, even if they disagree with yours
- Professional
  - Every professional engineering organization adheres to some form of ethics
- Institutional
  - Often a source of conflict with the personal and professional levels
- Macroethical (ethics applied to complex, large-scale anthropogenic systems where outcomes are highly uncertain)
  - The level of ethics necessary for sustainable engineering in the Anthropogenic world

FIGURE 12.1 Sustainable Engineering: Levels of Ethics. This figure presents the four levels of ethics that the sustainable engineer should be familiar with. In most cases, the first three will be apparent, although they may not be recognized as different levels; the fourth level, however, is one that is just beginning to be recognized as critical.

Although there are a number of ways to parse a discussion on ethics, let's begin by breaking ethical systems into four levels (Figure 12.1).

At the most obvious level, even if he or she has never had any formal training in ethics, every engineer operates within a relatively well-defined personal ethical system. These may differ significantly in detail depending on one's cultural background, but a number of general principles appear to be fairly widespread. A recent survey found significant shared values across cultures, countries, and religious affiliations, for example, listing eight ethical principles common to all of them (Box 12.1 discusses a formulation that extends across many belief systems in different forms, the Golden Rule):

1. Love: compassion, caring, empathy
2. Truth: honesty, integrity
3. Freedom
4. Unity: solidarity and community
5. Tolerance: diversity, minority rights
6. Fairness: equity, justice
7. Responsibility: accountability
8. Respect: for people, for environment, for diversity of species

BOX 12.1 Respect and Reciprocity: The Golden Rule

Discussions of ethics often reduce to either of two oversimplistic poles: an assertion that one particular set of ethics is clearly superior (because the proponent happens to accept that belief system), or a contrary, and unhelpful, embrace of relativism (no ethical system is preferable to any other; anything goes—the obvious response to which is that regardless of one's position, there is something evil about, for example, Nazi ethics). In fact, there is substantial evidence that may indicate ethical principles are fairly universal, with the important caveats that different cultures will

express them differently, and that different cultures will prioritize ethical principles differently. Thus, for example, Chinese culture, because of its deep appreciation of Confucian ethics, tends to emphasize harmony and hierarchical duty; American culture, with its sense of the importance of the individual, tends to value, for example, free speech more. There is both a universality and a particularity to ethical frameworks that should be appreciated by the sustainable engineer.

One of the key values that many cultures emphasize is the combination of respect and reciprocity that, in Christian theology, is known as the Golden Rule. But in fact that general value is widespread among world belief systems:

> ***Confucianism, Sayings 15:23:*** "What you yourself do not want, do not do to another person."
>
> ***Islam, Forty Hadith of an-Nawawi, 13:*** "None of you is a believer as long as he does not wish his brother what he wishes himself."
>
> ***Buddhism, Samyutta Nikaya V, 353.35–342.2:*** "A state which is not pleasant or enjoyable for me will also not be so for him; and how can I impose on another a state which is not pleasant or enjoyable for me?"
>
> ***Hinduism, Mahabharata XIII 114,8:*** "One should not behave towards others in a way which is unpleasant for oneself: that is the essence of morality."
>
> ***Jainism, Sutrakritanga I, 11,33:*** "Human beings should be indifferent to worldly things and treat all creatures in the world as they would want to be treated themselves."
>
> ***Christianity, Matt.7.12; Luke 6.31:*** "Whatever you want people to do to you, do also to them."
>
> ***Judaism, Shabbat 31a, Rabbi Hillel:*** "Do not do to others what you would not want them to do to you."

What such surveys show is not that all cultures have the same values, or different ones, for that matter, but that certain modes of behavior are regarded as important and widespread. Although one would hesitate to say that any value is universal, it is clear that some, such as respect for others and reciprocity, have a value for social cohesion and group function that would make it surprising if they were not widely adopted. As an additional observation, it is also interesting that these values—respect for others, and treating them decently, as you might wish to be treated in their place—are also important values for anyone, such as a sustainable engineer, who seeks to solve difficult problems in the real world without resort to unilateral force.

It is important to note, however, that cultures often prioritize ethical principles differently: thus, for example, the American culture tends to value highly individuality and individual success, while Confucian cultures, such as China tend to value highly community and obedience based on hierarchy and familial or social position. It is not that Asians can't succeed as individuals—

as Japan, China and Korea have demonstrated throughout history—nor that Americans are completely uninterested in community. But they prioritize their values differently. It is also important to note that while cultures might share the same basic value—say, freedom—they might well differ in how they define the principle in practice. Cultures that discriminate against women, for example, occasionally justify it by claiming that it frees women to be themselves. Cultures that practiced or still practice slavery may define freedom as accruing to persons, but then define their slaves as "not persons."

In addition to one's personal ethics, any practicing profession will have certain ethical standards to which its members are expected to adhere. These ethical standards are absorbed in part through the educational process. Engineering students do not just learn the substance of their discipline; they at least indirectly also learn about the behaviors, such as plagiarism or falsifying data, that are considered unethical and that are usually punished if detected (see Box 1.4). This informal but powerful professional ethical structure is augmented by formal ethical systems. Engineers, for example, have long had explicit codes of ethics applicable to individual professional behavior. Most fundamentally, such codes of ethics emphasize the responsibility of the engineer to society. Figure 12.2

- Preamble

  Engineering is an important and learned profession. As members of this profession, engineers are expected to exhibit the highest standards of honesty and integrity. Engineering has a direct and vital impact on the quality of life for all people. Accordingly, the services provided by engineers require honest, impartiality, fairness, and equity, and must be dedicated to the protection of the public health, safety, and welfare. Engineers must perform under a standard of professional behavior that requires adherence to the highest principles of ethical conduct.

I. Fundamental Canons

Engineers, in the fulfillment of their professional duties, shall:

- Hold paramount the safety, health, and welfare of the public.
- Perform services only in areas of their competence.
- Issue public statements only in an objective and truthful manner.
- Act for each employer or client as faithful agents or trustees.
- Avoid deceptive acts.
- Conduct themselves honorably, responsibly, ethically, and lawfully so as to enhance the honor, reputation, and usefulness of the profession.

FIGURE 12.2 National Society of Professional Engineers Code of Ethics. This figure displays the preamble and fundamental cannons of the National Society of Professional Engineers Code of Ethics; other professional engineering societies have similar codes of ethics in their governance documents. Although these are usually general, they establish baselines that engineers are expected to follow; professional and even legal penalties may result if they are not.

(Source: National Society of Professional Engineers (NSPE). "NSPE Code of Ethics for Engineers." http://www.nspe.org/Ethics/CodeofEthics/index.html: accessed January, 2010. Reprinted with permission.)

provides an example of the Preamble and Code from the National Society of Professional Engineers; other professional engineering societies such as the IEEE or Institute of Civil Engineers (ICE, in the United Kingdom) have similar provisions. These principles or canons are generally expanded by examples and rules of practice that further explain how they are applied in practice. Violation of such codes may entail penalties ranging from expulsion from the professional organization to withdrawing of professional licenses. More subtly, in legal cases in some jurisdictions failure to follow professional ethical standards can be evidence leading to civil (financial) liability or even criminal liability (for example, failure to follow professional rules of practice could be evidence of negligence in a court case).

But where do these ethical statements and codes come from? The brief answer is that ethical codes are based on a combination of experience, social norms, and ethical theories. These ethical theories, regardless of culture, tend to fall into two categories: they appeal for their validity either to some sort of rule set establishing preferable behavior regardless of circumstance ("That's the way we've always done it!"), or they appeal to outcomes of the act ("Because it protects the public, that's why!"). Figure 12.3 sets out the major theoretical frameworks that the engineer is liable to run into: the first, utilitarianism, is outcome based, whereas the other three are forms of rule-based ethics.

Utilitarianism generally holds that a choice is preferable if it provides the greatest good for the greatest number. Obviously, such an ethical system is compatible with democratic forms of government, in that majority rules is one indicator of greatest good for greatest number. For example, when a city decides to put a highway through an existing development in spite of vigorous opposition from a few families who would be displaced, it is making a utilitarian decision. In practice, of course, utilitarianism must be balanced by protection against extreme acts and majority prejudice. That many people

- Utilitarianism—what is ethical is that which produces the greatest good for the greatest number
- Duty ethics—what is ethical is to perform duties regardless of whether they lead to "good" outcome
- Rights ethics—mirror of duty ethics; people (or ecosystems, or living beings) have rights that cannot be violated
- Virtue ethics—actions reflecting good character traits are good; vices are bad; outcome of action is not relevant

FIGURE 12.3 General Ethical Theories. This figure shows the four main ethical theories that sustainable engineers should be familiar with. Of these, utilitarianism, an outcome-based ethical theory, tends to be dominant in many cases, in large part because engineering is an applied activity directed toward the solution of a particular problem in a way that benefits the client and the public.

benefited from slavery in the American South does not make it right because of the extreme damage inflicted on those who were made slaves. Utilitarianism is also very compatible with engineering, which as a problem-solving activity tends to produce results that are justified because they are valued by the client and the public.

Duty ethics, rights ethics, and virtue ethics are all ethical forms relying not on outcome, but on the desirability of the behavior, and usually on particular rules governing behavior. Thus, for example, the German philosopher Kant famously established as a core ethical requirement that people should be treated as ends, not as means (that is, should be respected for themselves, as moral agents, rather than treated as ways to accomplish something else). In its pure form, at least, this clearly breaks down quickly in a real society, where people—as employees or as soldiers, for example—are routinely treated as means. Similarly, rights ethics can get quite confused when, as some environmentalists argue, all living organisms are given rights. If everything has equal rights, then at some point no one has rights. (The most difficult ethical issues tend to arise not when one is faced with good versus evil, but when one is faced with two options, both of which are justifiable, and one of which must be subordinated to the other.) Virtue ethics can be similarly challenged. The question is not so much that some people are obviously more virtuous than others, it's who gets to define what virtue means in the first place. For example, some religious groups are quite adamant that they, and they alone, know what virtue is, so that everyone who behaves in a manner that they disapprove of is unethical. This obviously can lead to conflict in a heterogeneous community. The difficulty with rule-based ethical systems such as these tends to arise not when one knows exactly which rule to apply, but when the situation is more complicated than the rules allow for.

In a practical sense, many situations involve some sort of mixture of these theories. The requirement that engineers not lie about data, for example, may be considered virtue ethics, in that it doesn't say "unless the outcome justifies the lying because you think more people benefit from it." But the basis of that rule is utilitarian—it provides the greatest good for the greatest number when one can trust what engineers say about data, all things considered. Other professions make similar concessions to reality and practicality. The results of some cases in law are clearly unfair and wrong, and yet society tolerates those outcomes because, overall, an explicit and generally fair legal system is preferable to making up laws for each individual situation on the spot.

This points out an area of friction in many professional ethical systems, including engineering: that of scale. Because engineers work for larger organizations, private and public, they are not free to design or practice just as they please. Most obviously, economics and competition will affect almost any

engineering activity (put another way, you never have enough money to design just as you want). Thus, at a somewhat higher level, the idea of "social ethics" has been advanced to cover the important intersection between the practice of individual and professional ethics, and the institutional context within which those ethics are to be practiced or, in some cases, undermined. If you are told by your company to design a car that can be manufactured and sold for under $2,000 and weighs less than 2,000 pounds, and you are given a much shorter product design time frame than you're used to, you do your best to do so. If the gas tank then explodes in some cases, is it the fault of the engineering team, or the firm and management that insisted on the constraints even when suboptimal design decisions were a predictable result? (This is of course the Ford Pinto case.) If you create wonderful building designs, but you know they will be built by companies that, because of an endemic culture of corruption, will skimp on material quantity and quality, whose fault is that? Don't make the mistake of thinking that such things don't happen in developed countries. Such corruption is one reason the ceiling panel fell off the tunnel in Boston's Big Dig highway project, killing an individual.

But at least these kinds of dilemmas occur in, and are recognized as a part of, existing engineering practice in a market economy with recognizable laws and reasonably common cultural values. They are already in engineering ethics textbooks and used as examples in professional training. It is indeed the case that most of the ethical issues that a sustainable engineer is liable to run into never extend beyond these boundaries, and so can be adequately addressed at a personal or institutional level with existing ethical structures. For example, an engineer may resolve a difficult ethical issue by referring to her professional organization's ethical canons, or simply because she "knows" that a particular course is acceptable, or conversely, wrong. Figure 12.4 sets out a framework for performing such analyses: Determine what kind of situation one is in; determine what issues are factual in nature, which are definitional, and which are ethical; evaluate where the ethical responsibility lies; if it lies with you, review your options; finally, act appropriately. It is very important in real-world situations to perform a parallel analysis of what your legal responsibilities might be. What might seem on first blush to be the most ethical path could well create civil or even criminal liability in situations where, for example, you owe a contractual duty to your client, or the contract you're working under provides for specific remedies (e.g., if the contract says you will first discuss potential design problems with your client, do not begin by phoning the local news station).

This does not mean that ethical conundrums will not arise. They do regularly, which provides all the examples used in the engineering ethics textbooks. It does mean, however, that in general resolution of such conflicts occurs within a settled legal and cultural context. Another way of putting it is that ethical

- What kind of situation is it—personal, professional, institutional, macroethical, or mixed?
- Define the elements of the situation.
  - Factual: Some facts may not be known and some may be controversial (e.g., What is an acceptable chemical exposure level?)
  - Conceptual (e.g., What is a bribe and what is a gift? What is proprietary business information?)
  - Ethical: What ethical principles or codes apply?
- Understand conflicts, pressures, and power structure
- Evaluate where the ethical responsibility lies (With you? With your institution?), and to whom is it owed (Workers? Shareholders? Public?)
- Identify options and their cost/benefits (Negotiate? Quit? Blow whistle? Note whistle-blowers often lose. . . .)
- Act—not acting usually constitutes an (un)ethical choice in these cases.

FIGURE 12.4 Analyzing Ethical Situations. This figure illustrates one path for analyzing difficult ethical situations. Perhaps the most critical step is determining what type of situation one is dealing with, because often what appears to be an ethical problem is in fact a factual or conceptual issue. For example, obviously taking a bribe is unethical, but what is it when a politician who gives your firm a lot of business asks for a large contribution? In other words, the question may be as much factual—Is this offer a bribe under existing legal standards?—as it is ethical.

systems have evolved to operate at the same scale as human decision making and impacts—that is, at the individual, community, or institutional level. But in sustainable engineering, as you might suspect, this framework breaks down in some serious ways.

12.3 THE CHALLENGE OF MACROETHICS

As this book has illustrated in many ways, our understanding of the scale of human impacts is undergoing a fundamental change as we enter the age of the anthropogenic Earth. The climate change debates, for example, reflect not a temporary aberration in human affairs, but the fact that for the foreseeable future climate dynamics will continue to be shaped in significant ways by the activities of our species, and that the climate system and human economic activity now form an integrated system. What we have, therefore, is not a "climate problem" to be "solved"; it is a condition—the carbon cycle as integrated with human activities at global scale—that we have to learn how to live with and how to manage responsibly, ethically, and rationally. What is your responsibility as a sustainable engineer to the carbon cycle? The nitrogen cycle? The poor in India? Polar bears in the Arctic? Traditional codes of ethics and formulations of ethics provide little grounding here. We must begin from scratch.

We can begin by identifying three relevant characteristics of the Anthropocene that differentiate it from the traditional human systems within

which existing ethical structures have developed. The first is that the earth systems characteristics of the Anthropocene are neither "human" nor "natural," but are highly integrated composites of built, human, and natural subsystems. Second, and perhaps most important for ethical analysis, the human component of earth systems introduces reflexivity into them. What we think and learn feeds quickly and unpredictably back into the very system we're embedded in to begin with—thus the inherent unpredictability of human systems. Finally, earth systems are not just complex combinations of subsystems of very different types; they are also highly interconnected among themselves. After all, the attempt to manage global climate change is difficult precisely because the climate system is tightly coupled to human economic and technological systems, and to powerful cultural and ideological systems. It is also tightly coupled to other natural systems such as the carbon and nitrogen cycles; for example, growing significantly more corn for ethanol in an attempt to sequester carbon biologically also substantially increases the already profound and poorly understood anthropogenic changes in the nitrogen cycle. Because nitrogen is an important limiting factor in many ecosystems, changing nitrogen patterns has significant effects on biodiversity, especially in estuarine systems. And on and on. Each thread you pull on unravels something you hadn't thought of somewhere else. Again: This is not a "problem" with a "solution": it is engineering in the twenty-first century. Get used to it. More important, get used to thinking ethically and responsibly about it. But how?

As sustainable engineers, let's begin by thinking about technology systems. Consider, for example, the Portuguese caravel, a light sailing ship of the fourteenth century that was particularly adept at sailing into the wind and thus became the basic design enabling European global expansion. If the first caravel had sunk because of negligent design, ethical responsibility would have clearly fallen on the responsible ship engineer. It is highly likely that an annoyed king would have seen it that way, with unhappy consequences for the engineer. But it didn't sink, and the caravel, grown to galleon size over time and mounted with cannon using corned gunpowder, became the vehicle by which the new European nation-states established colonies, fought for control over oceanic trade, and established the lucrative, but infamous, slave trade. Few, however, would assign ethical responsibility for the effects of European colonization and the resultant Eurocentric global culture that the caravel enabled to that original ship engineer. There are simply too many unpredictable decisions and historical changes intervening between the design activity and the colonization outcomes (for example, China's decision in the 1430s to burn its competing fleet, and eschewing waterborne colonization, to turn inward). Or take another more modern example, the Internet. It's a complex, self-organizing system entirely constructed by humans, and yet its overall design and structure, not to mention its social and cultural implications, cannot be known a

priori by anyone. If a design team creates a new machine for the Net—a router, for example—and it bursts into flame every opportunity it gets, those engineers, and the firm that builds their design, would be held ethically and no doubt legally responsible. But if it turns out that virtual realities on the Internet become highly addictive to a subgroup of adolescent males, or social networking technology creates a different cognitive structure among young digital natives, few would assign ethical responsibility for such results to the design team; such effects are far beyond the capability of any engineer or design team to predict or foresee. The complexity of these systems is beyond the ability of existing human institutions, including ethical formulations, to address effectively.

Given the importance of sustainability to many people, it is not surprising that the sustainability community has made attempts to fill the gap between existing ethical formulations and the complexity of the systems with which sustainability in particular deals. Perhaps the most well known is the Precautionary Principle, which is especially popular in Europe. This principle basically states that new technologies should not be introduced until their effects are understood and the risk is confirmed as less compared to current practices. As stated in the UN's 1982 *World Charter for Nature* (http://www.un.org/documents/ga/res/37/a37r007.htm: accessed February 2011), for example, it commands that "where potential adverse effects are not fully understood, the activities should not proceed." Another sustainability example is the famous Aldo Leopold ecoethic: "A thing is right when it tends to preserve the integrity, stability and beauty of the biotic community. It is wrong otherwise." There are similar examples not in the sustainability realm, such as the political implementation of particular religious traditions, or efforts to simply extend existing ethical systems into broader regimes by, for example, making individual engineers researching technologies such as nanotech or biotech responsible for the behavior of the larger systems on which they work. Taken to an extreme, as some do, this becomes the equivalent of trying to make the ship designer of the caravel responsible for European colonialism.

These efforts are doomed to be unsuccessful because they fail to understand the difference between complex and simple systems, between the usual bounded ethical situation and the much different ethical challenges of dealing with complex earth systems. Thus, both the Precautionary Principle and the requirement that individual scientists be responsible for systems effects demand a knowledge of future paths, which is simply not possible. With any technology of more than trivial importance, you can never "fully understand potential adverse effects," because they are unpredictable and uncertain until they actually occur, and they don't occur until the technology rolls out across social, cultural, economic, and natural systems, all of which change in response. Accordingly, such proposals are either naïve or Luddite, amounting

to backdoor attempts to impede technological and scientific advance without appearing to. In contrast, most other approaches, such as the Leopold ecoethic, assume that a single value—in this case, the ecosystem—is adequate to provide ethical guidance in the Anthropocene. But this critical assumption is also mistaken. The behavior of these systems at scale is of an order of complexity that cannot be captured by any single worldview. Moreover, the related assumption that ethical decisions about earth systems by particular interest groups based only on their values (and ignoring or downplaying conflicting values that others may hold) will allow them to determine the paths of global technological systems is demonstrated by recent events to be wishful thinking. The strong European Union reaction against GMOs has not prevented the United States, China, Brazil, Australia, and other countries from rapid commercialization of the technology, and the Bush administration's efforts to halt biomedical research of certain types for religious reasons did not prevent Korea, Singapore, and others—indeed, did not prevent California—from continuing the rapid advance of the science.

What, then, is required for a reasonable macroethical system—that is, an ethical system operating at the scale of earth systems? First, macroethical frameworks must be able to accept, respect, and work with mutually exclusive but equally valid worldviews. It you're in charge of designing the Everglades, you must be able to understand and work with environmentalists, deep greens, developers, farmers, scientists, Native Americans—with many different groups with many different perspectives and priorities. Especially given the global scale of technology systems, multiculturalism is not simply political correctness, but is the only way many projects and designs can be successfully planned and implemented.

Second, macroethical systems cannot rely on assumptions about outcomes, because the future behavior of complex adaptive systems is highly uncertain. Judging actions before or while they are being taken is possible with simple systems where outcomes are reasonably foreseeable; it is not possible where systems responses cannot be known in advance. This means that the utilitarian calculus that often underlies scientific and technological policies is inadequate in the case of these large and complex systems, for the simple reason that utility cannot adequately be known beforehand. This has the important implication that, in practice, macroethics requires real-time dialog with the systems in question, rather than one-off snapshots—and it requires real-time adjustment of policies, and even ethical assumptions, as the systems change in the real world.

Third, many of the systems of interest—the climate system, carbon cycle, or designed biodiversity, for example—will extend long enough so that the cultural models and assumptions that ethical systems are built on will themselves evolve and change. This is a serious challenge to our current understanding of

ethics, where we tend to assume that different situations may require different ethical approaches, but the underlying ethical principles themselves don't change. (In fact, because acknowledging that ethical systems change undermines their claim to fundamental and unchanging validity, any recognition of change in ethical parameters, or effort to be flexible, will likely be strongly resisted.) Thus, a macroethical system should focus on process rather than individual outcome, substituting a continuing dialog rather than an "ethical solution." It also must be able to function in an environment where norms, ethical principles, and worldviews are understood to be contingent. In this, macroethics reflects the point made previously that many earth systems challenges, such as climate change, are not "problems" that can be "solved," but conditions arising from the dawn of the Anthropocene that we must learn to manage systematically and continually.

Ethics at this level of flexibility is a new and, for many people, challenging concept, but remember that it is the nature of the systems we have created and are affecting, not just whimsy, that imposes these requirements. It is also important to remember that flexibility is not necessarily the same as relativism—that is, a situation where "anything goes" or ethics aren't relevant. Rules that are contingent when the system is taken as a whole may well be valid within bounded subsystems that remain locally stable. Designing a facility to be used in a particular culture will still need to obey the rules of the culture, even if globally a more flexible design is appropriate. McDonald's in India does not serve hamburger, because the cow is sacred in Hindu cultures. The local rule is, "Don't insult your host culture by serving beef," while the global rule is more contingent: "Serve what the local culture likes and considers appropriate and ethical." The technology in Japanese mobile phones to access the Internet differs from that in American mobiles, because the two cultures differ in their use of personal computers, but in both cases the rule for ICT is the same: enable high-quality, robust access to Internet functionality.

This may all sound somewhat vague: What, exactly, is the sustainable engineer to do about all this systemic ethical contingency? Although the current lack of theory and institutional structure regarding macroethics is apparent, that does not mean that engineers and technologists (and others) cannot take pragmatic steps now to support the development of a macroethical capability suitable for the Anthropocene. To begin with, although it is apparent that it is inappropriate to charge individuals with knowledge that they cannot hope to have about future developments, it is possible to impose an individual ethical responsibility to ensure that mechanisms are established by which technical communities and society at large can dialog with complex adaptive systems.

In determining what such a mechanism might look like, one begins with the observation that the challenges of maintaining an open-ended and

long-term macroethical dialog require an institutional, rather than an individual, host. Given this insight, rather than focus on individuals as the locus of macroethical implementation, engineering and other professional societies would seem to be good initial candidates for such a responsibility. Accordingly, each of us as engineers, technologists, or managers may not bear individual responsibility for the scientific and technological systems that surround us and that we contribute to, but it does seem incumbent on us to accept our ethical responsibility to push our professional organizations, such as the IEEE, the ASCE, the AAAS, and the national academies of engineering that most countries have, to create institutional frameworks within which an ongoing macroethical capability is established. In so doing, we may begin to develop ethical frameworks that remain based on individual initiative, choice, and ethical responsibility, but that also begin to reflect the increasing complexity of the problems, options, and constraints that characterize the anthropogenic Earth.

Exercises

12.1 Roger was employed as a software programmer in a company that designed robots. James, who worked in an assembly line that used such robots, was severely injured. If Roger had made a simple calculation error, should it also be considered an ethical violation? How would different ethical frameworks interpret his actions?

12.2 In 1978, William J. LeMessurier, a well-known structural engineer, discovered that the Citicorp Center in New York had been built with bolted joints by the contractor, even though he had recommended welded joints. He knew that these would cause critical structural failure under conditions of diagonal wind loading. He drew up new plans and ensured that welded plates were fitted onto 200 bolted joints. Should he be criticized for not getting the work done correctly the first time or for overstating the danger? Or should he be praised for recognizing his mistake and correcting it?

a. Write a memorandum outlining your first steps and what particular projects you think might be most valuable in helping your Initiative succeed.

b. Define a set of metrics by which you could determine whether the Initiative had succeeded or not.

12.3 You are a newly hired engineer for a large consulting firm that does many projects for your state. You are representing your company at a holiday party thrown by the governor, when the governor's aide comes up to you and asks if you want to make the usual large contribution to the governor's campaign fund. When you stall, the aide growls, "We'll remember if you don't, you know!"

a. What are the ethical problems involved in this scenario, and when do they arise?

b. What are the potential legal issues involved in this scenario, and when do they arise? Assume in your answer that under state law, bribery of a government official is illegal.

Annotated Bibliography

1. For the most up-to-date language and commentary on the ethical codes of professional organizations, it is best to check their Web sites (e.g., the National Society of Professional Engineers code of ethics can be found at http://www.nspe.org/Ethics/index.html). A more focused look at engineering ethics and emerging technologies is the National Academy of Engineering's 2004 volume, *Emerging Technologies and Ethical Issues in Engineering* (Washington, DC: National Academy Press). A broader reference source on applied ethics in general is H. LaFollette's 2003 volume, *The Oxford Handbook of Practical Ethics* (Oxford: Oxford University Press). Finally, an example of how real-time responses are required by complex systems, taken from the realm of policy and technology assessment, is the 2002 article by D. Guston and D. Sarewitz, "Real-Time Technology Assessment," *Technology in Society* 23, no. 4: 93–109.
2. A good introduction to the institutional level of ethical issues can be found in R. Devon's 2004 article, "Towards a Social Ethics of Technology: A Research Prospect," *Techne* 8, no. 1: 99–115. More generally, any engineering ethics textbook will provide examples that, although they may not be phrased this way, clearly illustrate how tensions arise between the individual engineer and the institutional constraints they face in various situations.
3. Although it was written in 1993, Mark Johnson's book on ethics as an emergent property of human cognition and culture, *Moral Imagination: Implications of Cognitive Science for Ethics* (Chicago: University of Chicago Press), remains an excellent read, and illustrates the power and promise of new and less dogmatic ways of framing ethics as a human institution.

CHAPTER

13 The Aral Sea, the Everglades, and Adaptive Management

13.1 INTRODUCTION: ADAPTIVE MANAGEMENT

Much of this book has discussed technology systems and their increasing integration in regional and global earth systems. But human impact on major earth systems has been around for a long time. For example, managing resource regimes—especially hydrological regimes associated with agriculture either directly or indirectly—has been a notable aspect of human evolution since agriculture first began. Even at small scale, such systems almost by definition integrate human and natural systems. On a regional scale, many people have argued that the need to manage resources as societies expanded and, especially, urbanized has been an important driver of civilization and institutional evolution. This is particularly apparent with water, a critical agricultural resource; many historians have documented the rise of "hydrologic civilizations" such as China, where, they argue, the need for design, construction, and maintenance of water systems at large scale may well have had a critical role in encouraging the evolution of a more complex civilization. Among other institutional innovations, management of regional water systems required a rudimentary bureaucracy, and methods of establishing and maintaining use patterns and technologies to modify natural water resources to better fit human needs. Certainly, hydrological regimes have from ancient times been a source of environmental change.

But agriculture is not the only driver for hydraulic engineering. More locally, large urban systems since antiquity have drawn on

surrounding watersheds for water and local rivers for sewage disposal. Ancient Rome, admittedly exceptional because of its size and the skill of its engineers, was at its height served by a system of fourteen aqueducts totaling 508 kilometers, which delivered a billion liters of spring water every twenty-four hours. It was a constant flow system, and the excess, along with the sewage of Rome, flowed untreated into the Tiber. As in China, except at much smaller scale, provision of this water created the need for a specialized bureaucracy. Led by the three *curators aquarum,* a staff of some 600 laborers, engineers, and plumbers was eventually required. In a real sense, the Romans were not just engineering water systems, but also regional resource systems, just as today New York City's water supply depends on a complex technological system that feeds water from much of the State of New York to the City (and on a complex bureaucracy, just as in Rome, and on a complex set of laws that control land use throughout New York State).

Until very recently, however, it was only the technological side of such systems, not the coupled restructuring of natural and human systems, which was considered by the engineers and the managers. Thus, such regional resource management regimes have generally been guided by relatively simple and anthropocentric goals and metrics: Does the system provide enough water for irrigation, industry, or drinking? Does it control flooding? Does it augment state power (often a major if unstated rationale for large dams and other "trophy" water projects)? Given the important economic and even survival values often at stake, the questioning of such projects has been infrequent. Even in today's world, rife with environmental and human rights NGOs, massive projects with significant environmental and social implications are often undertaken with little internal debate or conflict. Indeed, the Aral Sea example is exactly such a case.

But this one-sided perspective is less and less viable for a number of reasons. When there is a relatively stable population drawing on resources for local needs that are generally limited, as in the past, one may be able to safely ignore systemic effects of engineering resource regimes. But in a world with 7 billion people, and multiple demands on most regional and global systems, cycles, and resources, such an implicit boundary is no longer adequate. Moreover, as any engineer who has worked on a major project recently can tell you, the involvement of numerous stakeholders, often representing environmental and human rights values, is no longer occasional, but is a condition of modern engineering. Ask the engineers who worked on the Three Gorges Dam, or the Navajo Bridge, or any modern mining project, or chemical manufacturing facilities, or on cell phones for the African markets. In cases ranging from the management of forests and fisheries; to

the attempt to integrate economic development, environmental quality, and social values in unique regions such as those centered on the Black Sea, Baltic Sea, Florida and the Everglades, the Aral Sea, or the Great Lakes; to the implementation of new technologies such as fish farms, one finds far more scientific complexity and uncertainty, and many more conflicting values and powerful stakeholders, some local and some national and global. This complexity has, slowly and subtly, made traditional engineering approaches and assumptions, and concomitant management practices, obsolete; it is one reason why sustainable engineering is necessary.

In the category of regional ecological and resource regime engineering and management, which is a focus of much systems ecology, one response to this complexity has been the development of "adaptive management." Growing primarily out of the ecological literature, *adaptive management* is defined in one of the leading texts as providing "ways for active adaptation and learning in dealing with uncertainty in the management of complex regional ecosystems."[1] In a more constrained definition, the General Accounting Office (GAO) describes adaptive management in a report on the Florida Everglades restoration project as "an iterative approach for improving resource management that recognizes that because scientific information is imperfect and, as decisions are implemented based upon best available science, a structure must be in place to acquire better information and adjust the implementation actions accordingly to improve the probability of achieving the goals of restoration."[2]

The GAO definition is useful for several reasons. The first and most obvious reason is that it indicates that the theory of adaptive management is actually being experimented with in at least one case. It thus follows that this definition is based on field experience with a very difficult and complex system, the Everglades. On the other hand, a sustainable engineer might argue that it is in some ways too limited to the extent it implies that adaptive management responds only, or even primarily, to improvements in scientific understanding. In most cases, real engineering and management of these systems requires adapting not just to new scientific data and understanding, but also to shifting political, economic, and cultural patterns, as well as changes in the physical systems themselves, and changes in the system resulting from technology changes, not just improvements in the data and models derived from them. Adaptive management, in other words, is a technique for managing wicked complexity.

[1]Gunderson et al., 1995, at ix (see Annotated Bibliography).
[2]General Accounting Office, 2003, at 2 (see Annotated Bibliography).

As in the case of industrial ecology, adaptive management is a nice example of a still young field that can help inform further development of sustainable engineering principles and methodologies. Like industrial ecology, adaptive management is not just theory at this point, but is being developed in the context of actual challenges, such as re-engineering the Everglades and the Mississippi River Delta, restoring the vitality of the Baltic Sea, managing toxics in the North American Great Lakes region, managing fisheries in various areas around the world, and establishing long-term forestry management programs in New Brunswick and elsewhere. In that sense, adaptive management is being born of a somewhat experiential approach, with interesting parallels to recent developments in, for example, urban planning and corporate governance. In each case, the context within which institutional operations must occur has gotten far more complex, requiring different, less rigid, and more adaptable behavior on the part of the institution. This is suggestive of the process that sustainable engineering is also beginning to undergo, where practice and experience are not just useful, but critical in determining what works and what doesn't.

The following two cases should be read in that light. The first, involving the Aral Sea, was a relatively simple and ideological project, and partially because it went forward without much local consultation or input, essentially ignored the full complexity of the actual world. Accordingly, it has had many unfortunate environmental and social impacts. Essentially, for reasons primarily involving the technocratic emphasis of marxist ideology and culture, and an economic, production-oriented focus, there were no adaptive management techniques used in this case. The second, the design of the Florida Everglades (and by extension, much of Florida), is a current effort that displays many of the principles of adaptive management—and, it must be said, the difficulty of implementing them in the real world. Both are similar in that they are hydrological design projects at regional scale and necessarily involve deep couplings among human and natural systems; both cases also involve, serious conflicts between desirable ends, such as economic development, environmental quality, and biodiversity. The approaches taken, however, differ enormously, as do the results to date.

13.2 THE ARAL SEA

The Aral Sea is actually a lake located in central Asia, an ancient center of civilization and trade now including Tajikistan, Uzbekistan, and a large part of Turkmenistan, Kyrgyzstan, and Kazakhstan. At the beginning of the

twentieth century, the population in the area was approximately 8 million, primarily supported by agriculture that in turn depended heavily on irrigation: About 3.5 million hectares were irrigated at the time. This reflected historical patterns, as irrigated agriculture in the region dates back to approximately the sixth century and has always been a feature of human settlement in the area. The population in the region is now well over 50 million people. More important, over the last century, especially the last thirty-five years or so, the area irrigated has grown to over 8 million hectares, because of the diversion of 94 percent of the flow of two rivers, the Amu Dar'ya and the Syr Dar'ya, primarily to grow cotton; these rivers both fed the Aral Sea before being diverted. This was a major project of the Soviet Union; cotton had long been favored by Soviet central planners because it could be exported for cash. (It was known as "white gold," and the Soviet Union became a net exporter of the crop as early as 1937.) Unfortunately, as in many cases in the Soviet Union, application of technology in fields not associated with national security was rough, ready, and inefficient. Thus, the irrigation canals were inefficient in terms of water transport; some estimates are that only 30 percent of the diverted water, carried in unlined canals through sandy desert soils, reached its destination. Moreover, the cotton that was grown was of such a low grade that it frequently could not even be sold on competitive world markets.

The results of this reengineering of regional hydrologic resources were dramatic. The United Nations Food and Agriculture Organization estimated that historical inflow to the Aral Sea was some 56 cubic kilometers per year, which shrank to about 47 in the period 1966–1970, collapsing to 2 by 1985. As a result, the Sea, which only decades ago was the fourth largest lake in the world, in a relatively few short years lost about half of its area and some three-fourths of its volume, with concomitant drops in sea level and increases in salinity. (Figure 13.1 illustrates the historical and current extent of the lake.) Ecological, soil, and hydrological impacts were exacerbated because the shift to intensive irrigation and river exploitation was accompanied by significant drainage of the river deltas and their conversion to agriculture.

In addition, substantial new areas of desert have been created; the resulting Ak-kum desert did not even exist thirty-five years ago. Other impacts, arising both directly from reduction of the Aral Sea and indirectly from land use practices (e.g., salinization of agricultural soils), are regional in scope. By some estimates, some 40 to 150 millions of tons of toxic dust per year are currently generated and spread primarily by wind, with substantial detrimental impacts on regional agriculture and human health. Blowing dust has also resulted in significant melting of glaciers in the Pamir-Altay, Saili Alatau, and Akshirak

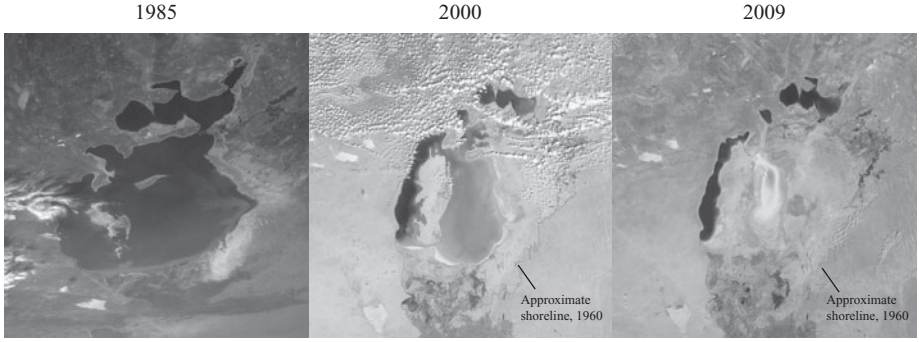


FIGURE 13.1 Shrinkage of the Aral Sea from 1985 to 2009. Note the outline of the Sea identified in the latter two images, which indicate that even the 1985 boundaries are attenuated from historical limits.

(Source: NASA.)

areas. More fundamentally, perhaps, the substantial change in such a large geophysical feature has modified regional climates, with potential impacts on the climate regimes of China, India, and southeastern Europe. In the region itself, the summers are shorter and drier, and winters longer and colder, with reduced vegetative seasons and significant changes in precipitation patterns. Degrading environmental conditions also took their toll on indigenous species. The number of bird nesting species dropped from over 170 in earlier years to around 35 today. Increasing salinity led to the loss of 20 out of 24 indigenous fish species.

Changes of this magnitude of course were coupled with economic and social impacts. Thus, a vibrant local commercial fishery saw a drop in the fish catch from 44,000 tons in the 1950s to zero today, with a concomitant loss of 60,000 jobs. The economic collapse implied by the disappearance of the major local employment opportunity exacerbated the health impacts that toxic dust and degraded environmental conditions were already imposing on local communities. But there were more unusual effects as well, including serious concern about the possibility of the release of biological warfare agents previously contained because they were worked with on an island in the Sea (Voskreseniye Island) which is now becoming a peninsula.

What should the sustainable engineer take from this case study? The most obvious initial observation is that this case study is not atypical. Similar water management activities have been going on for many centuries, often accompanied by unanticipated or undesirable impacts: soil salinization, susceptibility to drought, impacts on other (especially

aquatic) species (the latter generally not a concern until very recently). But hydrological engineering at all scales is perhaps as archetypal of the human species as urban systems, and such projects often have significant political implications. Indeed, in ancient China massive irrigation projects were a signature of, and a major coevolved institution supporting, the development of highly organized feudal power as long ago as the Chhin state in the Wei valley, some 2300 years ago. Over historical time, a number of societies are believed to have been weakened, if not collapsed, because of their failure to manage water resources, primarily for agricultural uses. Even today, by far the largest use of water resources is for agricultural irrigation.

But few cases lead to the train wreck that is the Aral Sea. (The few similar cases that come to mind, such as the collapse of the Iraqi Marshes, and the destruction of the Marsh Arab societies that lived there, often involve deliberate attempts to destroy the region. The Aral Sea case appears to have been inadvertent.) This leads to a second point that by now should be increasingly familiar to the student of sustainable engineering: Such projects necessarily reflect historical, ideological, and political contexts. In this case, for example, the background included a number of historical peculiarities. One was the Soviet fixation on economic development at all costs, and the underlying decision to exploit agriculture to create surpluses for military and industrial development. Such policies obviously favored any projects that would facilitate production of export-oriented crops such as cotton. More fundamental and more important, perhaps, was the Soviet Marxist predilection for demonstration of political and even theological validity through large engineering works. Such massive state projects reflected, of course, not just the desire to show the Soviet Union as technologically advanced and competent, but also the necessity for rapid economic development if the Union of Soviet Socialist Republics (USSR) was to keep up with its Western competitors regardless of cost. The most compelling thing about the Aral Sea disaster, after all, is that it was planned.

But it was planned as a traditional big engineering project, not using the reflexive tools of adaptive management nor any of the principles of sustainable engineering. In this way, the Aral Sea is in many ways the opposite of the Everglades example that will be discussed next, especially because the changes in the resource regime in the Aral Sea case have been intentional. Thus, unlike the Everglades where causality has been complex and perhaps not well understood, the Aral Sea is an example of what can happen when modern scientific and technological knowledge and skill is tied to the political and cultural power of a state with little or no questioning of social and environmental impacts and implications. This is a very important point for sustainable engineering: It was not because the engineering carried out was

done badly under the circumstances. If one wants to implement this sort of regional agricultural development at the lowest price regardless of inefficiency of infrastructure, the design wasn't bad, considering that social and environmental implications were not part of the picture at all. It was the priorities of the state—the social level of engineering ethics—that led to the results. Such activities are less and less acceptable in today's world, even if the project engineering is technically competent, because stakeholders are unwilling to accept new Aral Sea scenarios in most countries now, and they are more able to at least raise relevant questions, even in fairly authoritarian regimes.

It is thus noteworthy that there were no countervailing political forces opposing the technocratic, absolute state and its belief structure. The case study of the Aral Sea can be seen as a classic example of unintended impacts, easily predictable in hindsight, emerging at a regional level. But in fact, "unintended impacts" is somewhat of a misnomer; "deliberate blindness" may be more apt. Although on the obvious level Soviet planners did not intend the destruction they wrought, the effects of such massive diversions of water inputs to a landlocked system are hardly difficult to discern. Thus, there was at another level an obvious deemphasis of any negative perturbations resulting from the project. Moreover, these effects were apparent over time and no mitigation efforts were undertaken, so it is not just a question of initial mischaracterization of system response, but a continuing unwillingness to implement any sort of adaptive management approach. The easy accusation in hindsight is deliberate mischief, but that oversimplifies reality. It would be interesting to know the extent to which absolute power, ideology, and lack of concern about, and input from, local communities contributed to a fundamental failure to even perceive what was happening. There was no consideration of any ongoing dialog with the relevant human and natural systems, or plans to change behaviors or policy on a continual basis. The sorts of avenues that in many democracies would have provided such input to central power structures—such as community groups or NGOs in Europe or the United States—simply didn't exist in the Soviet Union. Rigidity, not adaptability, was thus built into the system.

Another important observation from a sustainable engineering perspective is that in this case, as in many cases, there were benefits from the project that were apparent to the State and to the engineers. Like many large hydrologic projects—especially regional irrigation systems, canal and dam construction, and the like—what amounts to an engineered destruction of the Aral Sea resulted from efforts to support higher levels of agricultural activity. People who were growing cotton where they couldn't before clearly benefited. This is common across many sustainable engineering projects: The distribution of costs and benefits fall disproportionately on different communities. A client

may benefit; a local community may face increased externalities (costs that aren't considered as part of the economic calculus behind the engineering decisions). Fishing communities on the old Aral Sea lost; new agriculturists gained; the State felt it gained power and resources. An important part of sustainable engineering is trying to design such systems so that as they evolve over time, the costs can be reduced and benefits increased and, to the extent possible, costs and benefits can be loaded on the same decision maker. When this can be done, it is far easier to create designs and engineer systems that are more sustainable over a longer time.

As a final postscript, with the breakup of the Soviet Union there are new efforts to recover at least some of the Aral Sea. How successful these will be— they are at this point intentionally only partial—remains to be seen. But the Aral Sea as it was, and as its communities were, cannot be recreated at this point. Like any complex system, the Aral Sea region is path dependent; once you have changed the system, you can never go back to what was; you can only reconstruct something new that reflects some of what you valued in the old.

13.3 THE FLORIDA EVERGLADES

Southern Florida is an area of about 18,000 square miles characterized by agriculture, attractive living environments (with air conditioning), and industry. Population growth is robust. Although South Florida was the most remote area in the eastern United States at the time of the Civil War, with less than 8,000 in the whole area, some 6.5 million people live in the area as of 2002. Rock mining operations to build their homes and concomitant infrastructure is an important and politically powerful local industry with significant environmental impacts. And, of course, there is the Everglades, a vast and biologically diverse wetlands. Despite the tendency to romanticize the Everglades as ancient wilderness, this "river of grass" is only about 5,000 years old, has already gone through two very different regimes and is embarking on a third. The first was characterized by a broad and shallow flow south from Lake Okeechobee to Florida Bay and the Gulf of Mexico. The second, a result of substantial human modification beginning after the American Civil War (ca. 1850) and accelerating strongly in the first half of the twentieth century, resulted in shifting of drainage patterns from Lake Okeechobee to the east and west, significantly reducing flows through the Everglades. The third, reflecting a desire to reestablish substantial flows through the Everglades, is currently evolving. Figure 13.2 shows a satellite photograph of the area; the line of islands called "keys," that stretches from the bottom of Florida encompasses

FIGURE 13.2 Satellite Map of Southern Florida. Lake Okeechobee, the Everglades, and Florida Bay are clearly discernable.

(Source: NASA.)

Florida Bay, a shallow and highly productive estuary whose health depends heavily on the state of the Everglades (and thus must be considered as part of any Everglades engineering and management regime). Figure 13.3 provides a schematic of southern Florida's major zones; most of the development is located on the Atlantic Coastal Ridge on the east coast, and along the west coast.

Much of the initial political and engineering structure of the Everglades dates from the 1930s. On May 30, 1934, Congress authorized the Everglades National Park, which was dedicated by President Truman in 1947; it was the first national park to be established because of its biological values. In 1938, the Hoover Dike, which flanked three quarters of Lake Okeechobee with a berm extending eighteen to twenty-two feet above the normal lake level, was completed. Such changes have over time resulted in

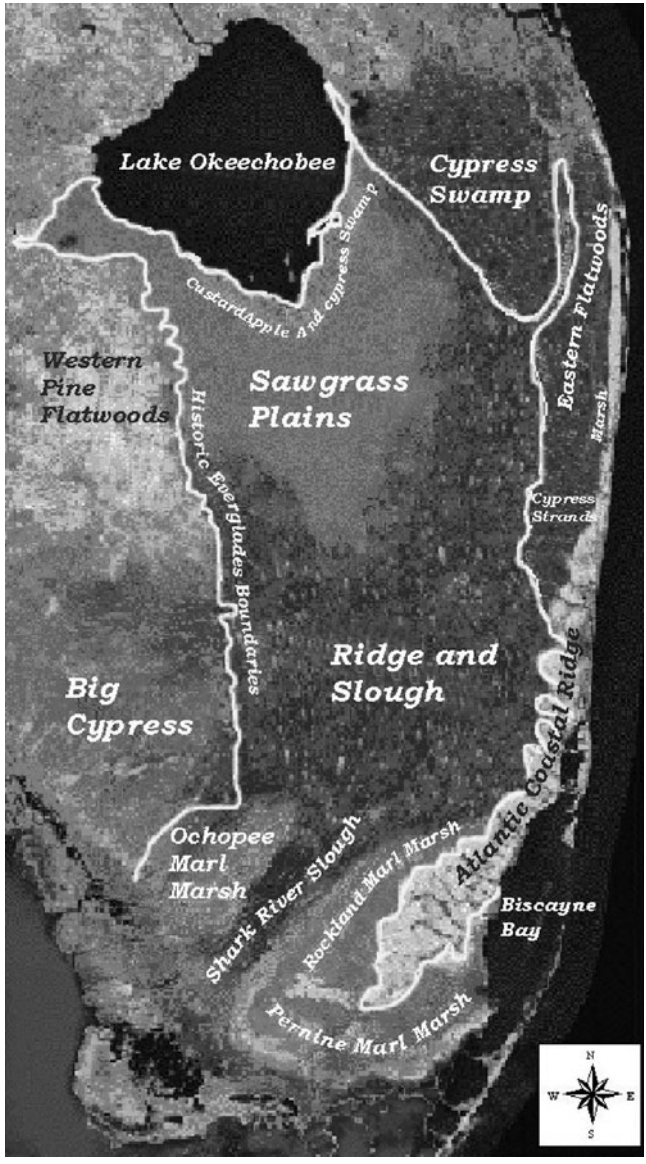


FIGURE 13.3 Major Zones of the Everglades. Most of the population lives along the Atlantic and Gulf of Mexico coasts.

(Source: U.S. Geological Survey.)

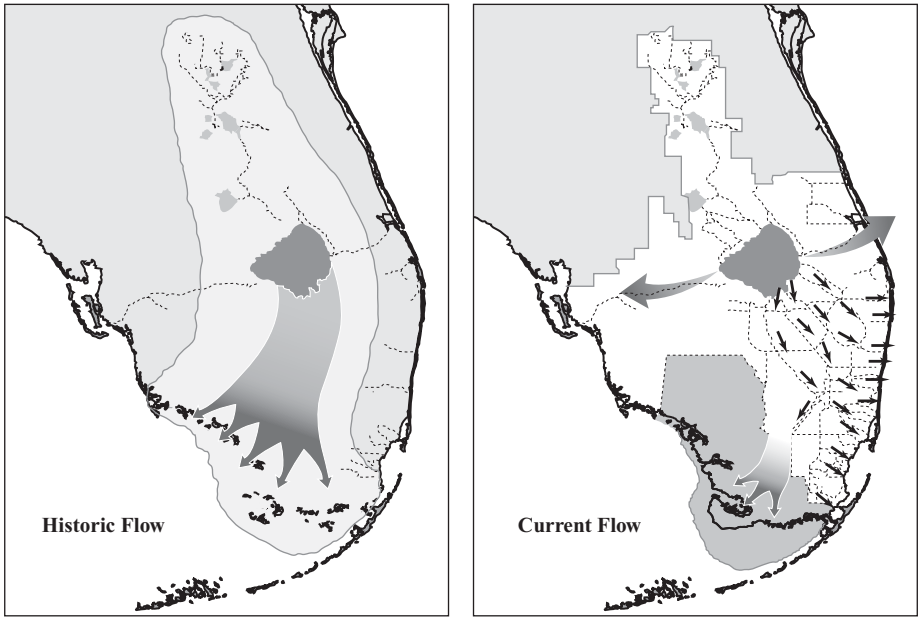


FIGURE 13.4 Florida Everglades: Historical and Modern Flow of Water. In southern Florida, the extent to which water that used to flow through the Everglades to Florida Bay has been diverted is apparent.

(Source: NASA.)

very different flow patterns in the Everglades (Figure 13.4), which naturally have had significant impacts on virtually all other aspects of life, human and biological, in South Florida. In particular, the "cultural cognitive dissonance" between preserving the biology of the Everglades and economic and agricultural development has consistently characterized the modern history of the Everglades. Development and technology state have, of course, long been linked, but the role of technology in shaping the Everglades is not just the obvious, such as hydraulic engineering for flood control, agriculture, and urbanization. It is also subtle. For example, a critical technology shaping southern Florida has been air-conditioning for both homes (invented in 1902 by Carrier) and cars (first used in the 1939 Packard, with mass market availability as an option in the 1954 Nash), without which highly uncomfortable hot and humid summers would discourage many potential immigrants.

The engineering of the Everglades in the twentieth century affected not just water quantity, but water quality. It is common knowledge that this unique ecosystem has been altered by an 1,800-mile network of canals built primarily over the last half century to support agricultural and settlement activity, which

has diverted some 1.7 billion gallons a day of water flow. Equally important, however, is the change in water quality that has occurred over that period. The original water of the Everglades was extremely low in the nutrient phosphorous, much lower than current flows, and increases in phosphorous have played a significant role in the shift in the dominant vegetation in the Everglades from sawgrass, which requires low phosphorous levels, to cattails. Increasing levels of phosphorous also changed other, less obvious, components of the ecosystem, including degradation of the periphyton and the food webs that depend on it. Concomitantly, invasive species, especially the melaleuca, Old World climbing ferns, Brazilian pepper, and Australian pine, are outcompeting native species, aided by human disturbances such as draining marshes and more frequent fires. The nesting success of birds, a predominant animal form in the Everglades, has declined by over 90 percent since the mid-1930s according to a number of estimates. Changes in the Everglades have, in turn, impacted Florida Bay, through perturbations in salinity levels, nitrogen cycling, and other mechanisms, leading to die-offs in important species such as the sea grasses.

The natural cycles that once defined and supported the Everglades in a stable condition, including rainfall and water distribution patterns, and nutrient cycles, have been profoundly affected not just by human settlement patterns, agriculture, tourism, industry, and transportation systems, but also by the various management regimes attempted over the past 100 years. Indeed, the increasingly powerful coupling among the natural, built, and human systems that interact in this sort of earth system is well illustrated in this case. Consider, for example, perhaps the most important single factor in the biological structure of the Everglades: the sugar subsidy that supports the predominant form of agriculture in the area. There are almost half a million acres of sugar fields between Lake Okeechobee and the Everglades, which exist because of a U.S. government subsidy to sugar producers that runs over a billion dollars a year. But sugar, and more broadly agriculture, are only one of a number of competing interests; human settlement, with concomitant demand for water and impact on water quality, and extensive mining activities, primarily for limestone for building, have also had significant impacts on drainage patterns, and thus biological system structure, in South Florida.

In response to the imminent danger of collapse of the Everglades as a functioning ecosystem, a $7.8 billion Everglades "restoration" project was begun, initiated by Section 528 of the Water Resources Development Act of 1996 (P.L. 104-303), which established the intergovernmental South Florida Ecosystem Restoration Task Force. Its intent is to restore water quality and water flow in nonhuman systems to functional levels, while continuing to support industrial, agricultural, settlement, and other human activity, and

avoiding flooding and other systems responses deemed unacceptable. The hope is to design and preserve a desired ecological community while maintaining human activity at a politically acceptable level, a challenge that will require implementation of over 200 separate projects. An important part of the plan is its emphasis on continuing dialog with both the human and the natural systems that, integrated together, compose the project, including continuing scientific review of both plans and the natural systems' reactions to what is already being done. Indeed, the 2003 U.S. General Accounting Office report on the overall restoration effort explicitly assumes adaptive management as the approach, even though it criticizes how well it has been implemented in practice. In addition, the focus of the adaptive management approach is on scientific issues associated with Everglades reengineering and to some extent the technological; the other dimensions of adaptive management, such as stakeholder involvement and political adjustments, are less effectively addressed.

Whether the various, somewhat conflicting, goals of the Everglades restoration effort can be met even partially is, to put it mildly, controversial. The process by which the restoration effort was tailored for approval at the state and federal levels inevitably led to compromises and tradeoffs. Not just environmentalists, but sugar growers, cattle ranchers, limestone mining firms, and development interests all participated in developing the plan, which thereby gained political acceptance. To some environmentalists, however, the compromises that resulted are unacceptable. It seems to some that the project relies too much on unproven science and engineering techniques (such as storing water in massive limestone lakes that are currently being quarried by rock mining companies for profit), favors development and industry over restoration and limits to growth, and, perhaps most difficult, will not improve water quality sufficiently (in particular, reducing phosphorous levels to the extent required to support indigenous species such as sawgrass). The complex adaptive system—with all its political, cultural, social, ideological, economic, technological, and natural domains—continues to evolve unpredictably.

13.4 THEMES

A few points are worth emphasizing from a sustainable engineering viewpoint. First and most fundamentally, it is obvious that the Everglades, and systems like it, are now, and will continue to be for the foreseeable future, products of human design and human choices. Massive human modification of the fundamental dynamics of the Everglades has already occurred. There is no "pristine" system to return to, and the Everglades will never be "natural" again. It will be a designed and engineered system, and it will display those characteristics—

including preservation of flora and fauna, if that is a design objective, and the system can be engineered to do so—that humans choose. The challenge is not to restore a hypothetical past state, but to exercise ethical and rational choice based on projections of systems evolution over time. Like the Aral Sea, like many fisheries and forestry examples, this case makes it very clear that in age of the anthropogenic Earth, the question is not whether natural systems are being engineered. Rather, it is whether we have the will and the honesty to admit that such is the case and assume responsibility for that which we are, in fact, already designing. As in other cases, the need for sustainable engineering does not arise because of choice or whimsy, but because of the world that we, as a technologically adept species, have already created.

The contrast between the Everglades and the Aral Sea is useful in part because it so clearly illustrates different approaches to sustainable engineering of earth systems. The Aral Sea example shows what happens when large regional reengineering projects escape the moderating influence of debate and discussion, and the cultural and ethical checks and balances that otherwise modulate policy dialogs in most societies. This does not necessarily equate to nondemocratic societies: the Chinese decision to build the Three Gorges Dam was certainly debated in that society (albeit sometimes not in public view), and the U.S. and U.K. decision to attack Iraq was in the opinion of many inadequately questioned in those societies. Lack of dialog, even at a somewhat simplistic level, should not be taken as a positive sign; rather, it is a warning sign because policy and sustainable engineering proposals at this scale are fragile and prone to failure because of unanticipated impacts. Moreover, note that whether each example is "sustainable" is not just a question of how long it lasts: with a strong central government able to override environmentalist and local concerns, the reengineering of the Aral Sea may last a longer time than the politically contentious effort to adaptively manage the Everglades. Clearly, the Everglades is a more robust example of sustainable engineering in practice, but that does not mean the system being designed and managed, the Everglades, is necessarily more stable over time; indeed, it may mean the opposite.

So what should sustainable engineering watch for? One danger sign is premature scaling up of pilot projects, sometimes as a result of market pressures. The geoengineering proposal to fertilize vast areas of South Pacific tropical reef areas with iron, thus reducing atmospheric CO_2 levels and dramatically increasing commercial fishery possibilities, at the cost of possibly destroying major reef communities, is one example. Another is the rapid scale up of corn-based ethanol as a gasoline additive; it is not that the technology itself is somehow invalid, but to ramp it up so quickly to such a high level of demand had distorting effects across political, economic, and agricultural systems around the world. Another danger sign is when specific

ideologies gain a disproportionate power to interfere in earth systems and are not adequately balanced by opposing cultural elements. This happened in the Aral Sea case, and some argue that it is happening in areas as diverse as responses to terrorism and climate change policy. For the sustainable engineer, the point is that public debate and dialog, although perhaps an annoyance, is helpful in defining problems and solution spaces, and in ensuring that an appropriate solution can be identified and, most important, is politically and culturally robust.

A final danger sign is when policy makers and the public fixate on only one part of a highly complex system, thus inevitably driving suboptimal and often dysfunctional solutions. An example of this is the current habit of trying to encapsulate disparate values in simple metrics such as the "carbon footprint." Although this simplifies public education about particular issues, it also drives policies, such as premature scale up of biofuel systems, which have significant costs in addition to their potential benefits. It is sometimes difficult to avoid such oversimplification in public discourse, but the sustainable engineer should not adopt such platitudes herself.

13.5 ADAPTIVE MANAGEMENT PRINCIPLES

Derived from the experience of natural resource economists, managers, and ecologists, adaptive management attempts to develop methods and tools that enable active adaptation and learning when dealing with the unpredictability and uncertainty inherent in the engineering and management of complex regional ecosystems. Because adaptive management is usually used with regard to ecosystems, and to protect environmental values, it is only a part of the competence required of sustainable engineering. Nonetheless, like industrial ecology and systems engineering, it offers a few general principles that are of use to sustainable engineering.

Adaptive management as a process does not emphasize institutional or procedural controls or mechanisms, but rather a different approach to understanding and interacting with complex regional systems. Because its focus is on "management," it does not produce specific engineering requirements, but the principles upon which it is based are equally applicable to those doing engineering and design work on complex systems such as regional resource regimes. Among its requirements:

1. Integrated policies, not piecemeal ones. This means that single-issue policy solutions should be avoided, or at best regarded with skepticism, and that policies that only address parts of the system are unlikely to be adequate in themselves. What adaptive management encourages is the development of policy regimes, rather than specific policies that react to single issues.

2. Flexible, adaptive policies, not rigid, locked-in ones. This is a reflection of the unpredictability of the underlying systems, which require a continuing dialog and adjustments in policy as they shift. This sort of flexibility is difficult to achieve in practice, in part because any meaningful policy tends to produce stakeholders who benefit from it and thus resist any changes in future, whether they are desirable or not. An example of this is the difficulty in implementing new environmental cleanup or control technologies, which may threaten existing economic interests or be viewed with extreme caution by environmental stakeholders.
3. Design, management, and planning for learning, not simply for economic or social product. Because the underlying systems are constantly changing and evolving, it is necessary to institutionalize learning as part of the engineering and management activity. In periods of slow change, knowledge can remain fairly stable over time; when systems are changing rapidly, continued learning is an important part of maintaining both the dialog with the system and the ability to adapt rationally and ethically to change.
4. Monitoring designed as a part of active interventions to achieve understanding and to identify appropriate response, not monitoring for monitoring's sake. With rapidly changing and unpredictable systems, it is important to know when they may be entering a particularly unstable or dangerous state. For example, many ecosystems can adjust fairly well to insult up to a certain point, when they rapidly and irreversibly collapse into another state. Metrics that can help identify such tipping points, and monitoring to determine when such tipping points may be approaching, are difficult to develop, but are a critical part of adaptive management (and sustainable engineering). In fact, part of the engineering of any project involving complex systems should be the development of appropriate metrics and the capability to monitor them.
5. Investment in eclectic science, not just in controlled science. This is a very difficult principle to implement in practice, because it cuts against both the reductionist perspective and the practice of peer review, that together form cornerstones of modern science and technology practice. But with complex systems, any coherent research program will necessarily be only partial and incomplete, and therefore potentially misleading. The challenge, of course, is understanding the difference between potentially useful research that may run counter to popular wisdom, and simply bad or mistaken research.
6. Citizen involvement and partnership to build "civic science," not public information programs to inform passively (or even worse, to keep people in the dark so they don't raise a fuss). This is an extension of the idea of dialog, and is particularly useful where the local community and civic

society have specialized knowledge (e.g., local foresters and fishermen). But it is also true more generally. Some engineers and scientists tend to believe that the public must share their ideas and values because their ideas are so rational and objectively self-evident, a position the public often does not appreciate. The more civic involvement there is, the more robust and effective a solution to a particular problem is likely to be.

Exercises

13.1 You are a recent sustainable engineering graduate who has just been hired to re-engineer the Aral Sea back to its original volume and productivity. What difficulties do you foresee in this assignment? Break your analysis down into four categories: technological, environmental, economic, and social.

13.2 You are the engineering aide to the governor of Florida, who is running for reelection and has asked you to write a memorandum to him on how to handle questions regarding the Everglades. Sensitive to the deep concern about this issue among environmentalists, developers, farmers and agriculturalists, workers in tourist industries, and industrialists, write a memorandum suggesting how the governor might want to address this issue using adaptive management principles.

Annotated Bibliography

1. There are a number of publications on adaptive management. Two of the better ones are edited volumes: the 1998 work edited by F. Berkes and C. Folke entitled *Linking Social and Ecological Systems: Management Practices and Social Mechanisms for Building Resilience* (Cambridge: Cambridge University Press); and the 1995 book edited by L. H. Gunderson, C. S. Holling, and S. S. Light entitled *Barriers & Bridges to the Renewal of Ecosystems and Institutions* (New York: Columbia University Press).
2. Similarly, there is a voluminous—and often highly contentious—literature on the Florida Everglades. Among useful and accessible sources are the 2003 U.S. General Accounting Office report, *South Florida Ecosystem Restoration: Task Force Needs to Improve Science Coordination to Increase the Likelihood of Success* (Report to Subcommittee on Interior and Related Agencies, Committee on Appropriations, House of Representatives, Washington, DC); and the Web site operated by the State University System of Florida, http://everglades.fiu.edu, accessed January 2010. For a glimpse of adaptive management in action, see the National Research Council's 2008 report, *Progress Toward Restoring the Everglades* (Washington, DC: National Academy Press).
3. The Aral Sea is also a popular topic. A good background Web site is that

operated by the United Nations Environment Programme and International Fund for the Aral Sea at http://www.unep.org/dewa/vitalwater/article115.html: accessed February 2011. An early (1998) article on the issue may be found at M. Feshbach, "Former Soviet Union's Environmental and Health Problems," in eds. B. R. Allenby, T. J. Gilmartin, and R. F. Lehman, III, *Environmental Threats and National Security: An International Challenge to Science and Technology.* (Livermore, CA: Lawrence Livermore National Laboratory, pp. 173–188).

CHAPTER

14 Earth Systems Engineering and Management: Sustainable Engineering at a Planetary Scale

Welcome to the practicalities of orchestrated planetary change.[1]

—R. H. Haynes

14.1 INTRODUCTION

So far the exploration of sustainable engineering has identified a number of critical themes: complexity, technological evolution (especially the Five Horsemen), the global scale of the integrated human/natural/built systems, and their emergent behaviors, that characterize the anthropogenic Earth. The need to integrate social, economic, technological, cultural, and environmental considerations into design and management of engineered systems—the essence of sustainable engineering in some important ways—is apparent. For the engineering student, already enmeshed in calculus, statics and dynamics, circuit design, and other necessary but difficult domains, this can seem an overwhelming challenge.

[1] R. H. Haynes, "Forward," in M. J. Fogg, *Terraforming: Engineering Planetary Environments* (Warrendale, PA: Society of Automotive Engineers, 1960, republished in 1995), p. v.

In some sense, then, a discussion of earth systems engineering and management (ESEM), which may be thought of as sustainable engineering at the planetary scale, might seem like piling on. In terms of complexity, it certainly is not trivial. But the student must remember two important things: What you are doing now, even in your most technical classes, is gaining an appreciation for the intellectual terrain that you will work in for the foreseeable future. It is neither complete nor static, and your most important goal should be to learn how to learn about it, not just become competent in the specifics. The latter will change; the former is what will help you adapt to those changes.

Second, for your professors, especially the more traditional ones, engineering was something they "learned" in school and subsequently practiced. For you, there is little if any difference between the two. If you do not constantly learn even as you practice, you will soon be rendered obsolete by the Five Horsemen and the rapidly changing social and environmental context of engineering. You are not, in other words, doing sustainable engineering because you want to. You are doing it because, if you don't, you will not be competent in the complex, rapidly changing world that humans have already created. You are doing it because you have to.

The Aral Sea and the Everglades have already been introduced as examples of regional scale resource regime ESEM. Other examples will help to flesh out the complications that arise at this level of sustainable engineering. For example, a number of scientists and engineers have proposed **geoengineering** solutions to global climate change; in general, such initiatives at scale would violate a number of ESEM principles such as reversibility and scalability, and seriously misunderstand the nature of powerful technology systems. In an analogous example, much urban design and planning in the 1950s and early 1960s involved imposition of technocratic views on cities without much consultation with, or even consideration of, the desires of citizens. When this was replaced by a postmodern community-based urban power structure, the situation didn't get much better. This is an important observation, because it suggests that it is not enough from a sustainable engineering perspective to get citizens engaged and involved, but that this must be done in productive and positive ways.

Fortunately, sustainable engineering at the ESEM level can draw on a number of areas of study and practice that are already fairly robust. Thus, the principles of industrial ecology, adaptive management, and systems engineering discussed in Chapters 7 and 13 especially are, properly generalized, certainly applicable to ESEM. Moreover, there is a growing literature on management and complexity, particularly in the corporate and industrial management area, which is quite useful, in part because both business managers and engineers face a similar type of challenge: creating solutions that work in

the real world for complex problems. In addition, sustainable engineering can draw on experience from a number of other fields of study, including urban and regional planning, sociology and the social construction literature, and the management of complex technological systems such as the Internet, global air transport and financial networks, military weapon systems, and the like to develop a set of initial, and obviously incomplete, ESEM principles. There is even a small but growing literature on terraforming, space architecture, and space colony design that may eventually be useful: after all, we may talk about terraforming other planets, but the one planet we have in fact already terraformed is the Earth. And, of course, although oversimplistic and naïve, the geoengineering debates are useful in deriving principles.

We will thus begin this chapter by briefly discussing the geoengineering and urban planning examples of ESEM (using more direct quotes than usual in the latter, so that you get a better feel for the flavor of that dialog). We will then turn to the principles themselves, with the understanding that they are preliminary and fairly general at this point, and should not be taken to represent a mature presentation of the field. Although any such set of principles is best understood as partial and advisory at this point, many of the principles have serious policy implications. Accordingly, the sustainable engineering student should approach the following discussion with caution (can we really include "all relevant stakeholders" in any serious policy discussion about regional or global ESEM issues that affect us broadly?), but also with the understanding that these principles can at least provide a framework for refining, and challenging, existing and proposed policies and alternatives.

14.2 GEOENGINEERING

At first blush, "geoengineering" sounds like a synonym for ESEM. It has, however, a very specific meaning in the climate change arena: It identifies technologies that are seen as capable of addressing most, if not all, of global warming by themselves. It is generally broken into two types of technologies. The first is "carbon capture and sequestration," or CCS, which in this instance includes such things as ambient air carbon capture using mechanical devices and some sort of chemical, such as sodium hydroxide (NaOH), that is able to latch on to carbon dioxide even at ambient atmospheric concentrations, around 350 to 390 parts per million. It also includes engineered biological systems, such as the idea of fertilizing the open ocean with iron, generally a limiting nutrient in such ecosystems, to encourage rapid phytoplanktonic growth (which absorbs carbon); the phytoplankton then die and sink to the bottom of the ocean, where they and the carbon they contain are sequestered. Another family of geoengineering technologies is captured under the general

title of Solar Radiation Management, or SRM. Techniques in this category include ones that involve placing something—clouds, balloons, space shades—between the Earth and the Sun, therefore reducing incoming energy and balancing the energy equation for the atmosphere. Clouds, for example, could be generated by injecting large amounts of sulfur aerosols into the atmosphere, thereby nucleating water droplet and thus cloud formation; alternatively, one proposal is to have ships sailing the ocean aerosolize ocean water, which contains salt particles that would also help create cloud cover. Clouds, of course, reflect sunlight and therefore are a means to increase the reflectivity (called "albedo") of the Earth.

Opposition to geoengineering by the climate change science and activist community has been strong and quite successful for a long time in preventing research funding or publication in the field. This reflects two main concerns. The first is the unquestioned risk of the technologies involved, which require direct and hardly subtle intervention in foundational planetary systems. (Proponents of geoengineering readily admit this point, but note that it is somewhat hypocritical, given that research that might help resolve such issues has by and large been successfully limited by the same community issuing the criticism.)

The second, which is far more important for many in the climate change community, is known as "moral hazard." This concern is that any consideration of alternatives except immediate deep cuts in carbon dioxide emission, especially geoengineering alternatives that appear to offer a means to control climate change, will reduce political and popular pressure on developed country consumers, particularly Americans, for reduced consumption and therefore reduced emissions. At one level, this is not a ridiculous concern, as it is indeed likely that changing consumption patterns, especially in the United States, would be far easier if climate change can be presented as both catastrophic and avoidable only with significant changes to current social and cultural patterns and values. But it hides a deeper and perhaps more troubling question: Is it appropriate and ethical for the climate change engineering and science community, by manipulating the kinds of research being done and the data presented to policy makers, to socially engineer society to reflect the values that they support? Certainly this contravenes relevant principles of ESEM because such an approach may provide fragile support for demanding deep economic and social change. If the stakeholders who are being socially engineered without being told find out about it, it is liable to lead to a strong backlash against the scientists and engineers involved, undermining the credibility of engineering and science just when it is most important to be able to trust the data.

But the sustainable engineer cannot take the behavior of the climate change community as determinative of the value of geoengineering or, for that matter, of the underlying reality of climate change, the physics of which are

quite clear. This means that a sustainable engineering analysis of geoengineering is not predetermined. Such an analysis raises several cautionary flags regarding geoengineering, at least at large scale.

Begin with the obvious. What we have here is a serious blind spot regarding technology. Sustainable engineering requires constant emphasis of the fundamental point that, if history teaches us anything, it is that any technology system potent enough to be interesting—the steam engine, railroad technology, the internal combustion engine, agriculture and hydrological engineering, and now social networking and wireless technologies—is potent enough to generate unpredictable and sometimes very fundamental changes in institutions, societies, cultures, other technologies, and economies. Moreover, it is clear that falling in love with any technology—be it corn-based ethanol, or the massive, information-dense military juggernaut that the United States unleashed on Iraq, or the various technologies that fall under the current rubric of geoengineering, or global circulation models—is very, very dangerous (Box 3.1). Geoengineering without serious attention runs the high probability of simply being a straightforward case of technoinfatuation: We cannot simply change fundamental systems with clumsy technologies and hope that we'll get away with it.

That said, although we must not fall in love with geoengineering technologies, it is also clearly the case that these techniques represent an important source of insurance in a future that is highly unpredictable and in which major systems—climate, ocean circulation patterns, and the like—are contingent in ways we don't really understand. Research on geoengineering should certainly be continued and indeed expanded significantly. The incremental benefits from the billions poured into climate science pale next to the importance of developing a sophisticated set of options from which we can choose. None of the technologies currently included in the geoengineering category are necessarily "bad," but they all are highly risky, especially if deployed at planetary scale. But the example of corn-based ethanol technology (Box 1.3) suggests alternatives. There's nothing wrong with the technology per se, and at small scale it's a perfectly good way of making ethanol. But scale it up rapidly to extremely large volumes, and it's a highly dysfunctional technology, causing political, economic, distributional, and social distortions that are entirely out of line with any putative benefits in terms of climate change or energy security. At limited scale, in other words, it's a useful and manageable part of sustainable engineering and the technological landscape; it is only when it is unexpectedly and rapidly scaled up that the effects become seriously nonlinear. Similarly, it is not that any of the proposed geoengineering techniques are "bad." A few reflective balloons in the stratosphere might be a good pilot project, and would indeed reduce albedo somewhat; a few ships spraying sea spray into the atmosphere to try to create cloud cover might be interesting; a little

iron fertilization of the ocean might be worthwhile. Among the important questions that must be answered, however, is how these technologies scale—and that can only be determined by careful research and pilot projects at a manageable, and reversible, scale. After all, one could have done the same with corn-based ethanol: scale up production until one noted the beginning of nonlinear effects (for example, farmers beginning to move serious acreage from other crops into corn, thus impacting food prices across the board and across the globe), then analyze the production level and see what the costs and benefits of further expansion of that technology might be. A crucial question, then, with all geoengineering technologies is when they operate in a linear, predictable mode, and when they go seriously nonlinear. Answering that requires research. Moreover, the "single-use" focus of geoengineering analysis is highly questionable and indeed dangerous: All technology systems have social, cultural, economic, technological, institutional, and psychological domains; focusing on only one (e.g., ability to reduce planetary albedo) is one of the major sources of weakness in the current discussions around geoengineering.

A sustainable engineering approach would also strongly suggest a need to expand the idea of what constitutes geoengineering. For example, recent work suggests the possibility that we'll be able to grow most of our meat in factories using stem cells in the not too distant future. This is in fact an interesting geoengineering technology. It would significantly reduce the impact of agriculture on climate (and other systems), because a cow is a fairly inefficient source of steak, and a pig of pork, all things considered. It also makes clear the systemic implications of powerful technology systems: the potential effects on existing agricultural economies, supply chains, land use, water use, and animal populations are apparent (not to mention the subtle but real psychological implications of changing food from a farm grown, to a factory, product). On the other side, the increasing possibility of a technology suite that supports radical life extension, at least among the elite, is generally unmarked in the climate change debates, although the implications of such a development would be profound (sort of "antigeoengineering"?). Any discussion of geoengineering is overly simplistic if it does not at least attempt to understand the implications of the Five Horsemen.

From a sustainable engineering viewpoint, then, the point is not whether we should do geoengineering. We already are—that's what anthropogenic climate change is. In that sense the arguments are already over: Geoengineering is here. The point is that the geoengineering debate and the technologies that it revolves around seriously need a dose of sustainable engineering. If the student wanted an example of why sustainable engineering is necessary and, indeed, critical, this case study provides a strong and immediate answer.

14.3 URBAN DESIGN AND MANAGEMENT, AND HIGH MODERNISM

Although the increasing urbanization of the human species was discussed earlier (Chapter 2), cities are also a useful way to demonstrate a continuing challenge to sustainable engineering: the danger of falling into high modernist hubris and technocratic blindness. **High modernism** is an engineering and policy approach often characteristic of technocratic elites (this is different from the way the term is used in literature and art, where it simply refers to the most developed aesthetic of modern work). It is generally characterized by an elitist and technocratic approach with a deep contempt for the "barbarism" of the past and the vagaries of the mere public, and it tends to ignore local context and culture in the interest of implementing "universal, scientific" principles, often at great human cost. In urban development and engineering, the high modernist approach is perhaps best personified by the work of Le Corbusier and Robert Moses. The former attempted to impose geometric perfection on the urban landscape, demanding that cities be planned with reference only to science and geometry. For example, Le Corbusier complained against "the sight of [the unplanned city's] . . . jagged, loose, diversified, multiplied and nerve-wracking spectacle; the sky is seen against a ragged outline and each house suggests, even by its very shape, some different order of thing. The eye is overwhelmed, tired and hurt. . . ." and that the city must be recreated as a totally planned, geometric, imposed order.[2] Similarly, Moses brooked no interference with his grand urban projects from the unwashed masses, famously declaring, "[w]hen you operate in an overbuilt metropolis, you have to hack your way with a meat ax. . . . I'm just going to keep right on building. You do the best you can to stop it."[3] His grandiose constructions across New York City, while they have inspired some, have been held by others to be responsible for destroying the human vitality of significant city spaces such as the Bronx. This was "high modernism," and in its time, it was the standard for urban engineering.

---

[2]Le Corbusier, [1929]1987, *The City of Tomorrow and Its Planning,* p. 71. The high modernist mentality that Le Corbusier brought to this task is well illustrated by his remarks later on in the same book regarding a construction project in the Alps (pp. 144–146): "The entrance to the high valley is by a steep pass. You hear a soft musical burden. It is the hum of well-oiled pulleys running along a steel cable . . . . the soft hum accompanies one everywhere; on waking at 5 a.m. in the hut of the Alpine club the ear catches the kindly sound which gives a sense of comfort, of security and order. . . . It may seem absurd, but one's mind goes back to the Giants building Valhalla. The gods are on earth: a touch on a lever in the machine-room; an organ plays softly over the wild landscape. . . . You say to yourself: Man is mighty; he assails Heaven itself. And in this Tower of Babel we hear the sounds of French, and the work goes forward. It is most moving, most captivating. The thing is magnificent!"

[3]Quoted in Berman, All That Is Solid Melts Into Air, 1982, p. 290.

Starting in the 1960s, critics such as Jane Jacobs began pointing out that this effort to impose Platonic idealism on a complex, dynamic, human system was ill-conceived, egotistical, inflexible, hubristic, and doomed to failure, a critique that has generally been accepted. Indeed, high modernism has generally been unsuccessful in urban environments, as anyone who has lived in a vibrant, loud, noisy, unpredictable, and confused city would have guessed. Thus another authority, Charles Jencks, in his book, *The Language of Post Modern Architecture*, commented in 1997 (p. 9):

> Happily, we can date the death of modern architecture to a precise moment in time. . . . it expired finally and completely in 1972, after having been flogged to death remorselessly for ten years by critics such as Jane Jacobs; and the fact that many so-called modern architects still go around practising a trade as if it were alive can be taken as one of the great curiosities of our age. . . . Modern Architecture died in St Louis, Missouri on July 15, 1972 at 3:32 p.m. (or thereabouts) when the infamous Pruitt-Igoe scheme, or rather several of its slab blocks, were given the final *coup de grace* by dynamite. [Pruitt-Igoe was a housing development consisting of huge rectangular slab blocks fourteen stories high, "constructed according to the most progressive ideals of CIAM (the Congress of International Modern Architects" and embodying all of the architectural rationality of Le Corbusier.]

But it was not just overly idealistic, overly simplistic built infrastructure that did in urban high modernism. It was also a fundamental shift in political power in urban systems from elites to communities and informal groups. From a political perspective, then, one might date the death of high modernism to the defeat of the planned Greenpoint-Williamsburg waste incinerator in the Brooklyn Navy Yard in New York City. Proposed in the early 1970s by the City as a successor to the huge Fresh Kills landfill on Staten Island, which was closing, the project was halted by an informal and shifting coalition of community and environmental activists in the 1990s, a result that would have been impossible in the Moses high modernist era. This incident is, of course, a nice illustration of a much broader trend, especially in the United Kingdom and the United States: Peter Hall in his magnificent 1998 book, *Cities in Civilization*, notes that this period (pp. 889–931):

> represented a change in *Zeitgeist*: not merely in Britain, but beginning in the United States in the closing stages of the Vietnam War, there had been a total inversion of established values: the entire concept of large-scale, top-down, professionally oriented planning was replaced by a belief in its opposite, bottom-up planning

> through local community groups, in which the planner was the servant of the public. . . . It coincided with the arrival of a new generation of activists who had come out of the schools and universities, many of whom were influenced by Marxism and were enthusiastically planning to make a series of local grass-roots revolutions which would challenge the hegemony—a favorite word—of the capitalist state.

This sea change was another nail in the coffin of traditional, "technology-only," urban design, and a precursor of sustainable engineering. Nonetheless, that doesn't mean that it wasn't problematic, or even dysfunctional. One of the unfortunate results, for example, was that the new "postmodern" environment created a political scale that was both inadequate to the physical scale of the systems involved (e.g., waste disposal), and reflecting a profound desire for the status quo, incapable of integrated future design (or as Hall put it, p. 897: "the ultimate myth was the basic 1970s belief in bottom-up community-based planning. . . it was a recipe for paralysis."). This should prove no surprise; political philosophers since Plato and Aristotle have written of the problems that arise when democracy slips into operational anarchy. What these case studies add to this general critique, however, is the additional complexity of scale. Community activism may simply fragment the political structure to a scale that is then unable to cope with the economic, physical, and temporal dimensions of the urban infrastructures and systems involved.

But such dynamics, although they must be perceived and understood, are not fatal to rational design and management, and thus to sustainable engineering. For example, by way of comparison to the Moses and high modernist approach, consider the approach taken by the engineers managing the Boston Central Artery/Tunnel (CA/T) project. The initial phase of their activities consisted of seeking public input and working with local neighborhoods and stakeholders, as they recognized that an infrastructure project of the scale proposed, which would change the structure of downtown Boston significantly, would fail unless it met the needs of local communities and cultures, as well as broader regional infrastructure demands. This process, with its cultural sensitivity and attention to nontraditional and different (even conflicting) stakeholders, is a significant contrast to the technological hubris that characterized high modernism, and has clearly become the predominant model over the past decades. It is, in other words, part and parcel of a sustainable engineering approach. But never forget that sensitivity to context and stakeholder involvement are necessary additions to, not replacements for, traditional engineering and management competencies, as the subsequent cost overruns, corruption, and engineering failures of the Boston CA/T project unfortunately illustrate.

However, the high modernist period, and cities themselves, also offer a suggestion of the limits to sustainable engineering, and the application of tools such as systems engineering, no matter how culturally and environmentally sensitive, to complex human/natural/built systems. In this regard, it is instructive to remember that efforts to apply systems engineering approaches and principles to urban redevelopment, which were especially popular in the 1960s, universally failed. Thus, systems engineering and management approaches can be applied *in the context of* complex urban environments *to specific projects.* But they are not sophisticated enough in themselves to engineer and manage urban areas as a "system of systems" because urban systems are of a higher level of complexity. Individual projects that may have important cultural dimensions are still bounded by the project definition and, accordingly, are less complex than the city system as a whole. The latter necessarily includes human and natural systems of many different kinds and at many different scales, with their inherent contingency and unpredictability.

14.4 THEORETICAL ESEM PRINCIPLES

The experience gained in many fields can be used to generate a set of general ESEM principles that provide some basis for current practice and, more important, for future development of more sophisticated ESEM approaches and methods. These principles will be grouped in three categories: theoretical, governance, and design and management. We will begin by exploring the theoretical principles, and turn to the others in following sections.

1. Given our current level of (often underappreciated) ignorance, perhaps the first principle should be precautionary: *Only intervene when necessary, and then only to the extent required, in complex systems.* This follows from the sustainable engineering principle that we need to treat all complex systems with respect and appreciate the extent of their unpredictability (and our ignorance). In that, it is somewhat similar to the medical admonition, "first, do no harm," and reflects the same rationale: When faced with a complex, uncertain, and unpredictable system, minimal interventions reduce the possibility and to some extent the potential scale of unforeseen and undesirable system responses. Note that this principle is very different from the Precautionary Principle (which essentially states that if one cannot know the risks posed by a new technology or practice, one should not introduce it; see discussion in Chapter 12). It is one thing to understand that minimal intervention produces the least risk of unanticipated system shifts; entirely another to stifle innovation unless an unrealistic knowledge of future states can be demonstrated. That said, the sustainable engineer must not fall into two potential traps. Because the Precautionary Principle is flawed does not mean that it isn't

regarded as desirable and important public policy by some stakeholders, especially in Europe, and it must accordingly be respected and integrated into design, engineering, and management decisions appropriately. In other words, even if it is not a robust basis for design choice in itself, it is an important component of the social context within which engineering is done. It indicates a deep skepticism about some technologies that is in itself part of the problem space. Second, the weaknesses of the Precautionary Principle do not in any way affect the responsibility to be cautious when intervening in complex systems, respectful of their ability to throw up completely unanticipated emergent behaviors and appreciative of the unpredictability of such systems. Precaution is a virtue, even if the Principle is naïve.

2. *Major shifts in technological systems should be evaluated before, rather than after, implementation of policies and initiatives designed to encourage them.* This principle sounds straightforward, but activists, the public, and policy makers frequently "fall in love" with particular technologies, making study of their potential downsides both economically and frequently socially infeasible. Even if you adore biofuels and electric cars (Box 3.1), it remains your responsibility to try to understand ways in which they may undermine social and environmental values if implemented improperly or scaled up too rapidly. Another example might be the proposal, popular among some, that agricultural rather than wood fiber should be used for making paper. That sort of technology shift should not be widely promoted until there is a reasonable understanding of the potential environmental, social, and economic implications. For example, one might want to ask how many sectors—energy, paper, wood fiber products—are converging on agricultural fiber and waste, and what the implications of such a scale up might be (in other words, are we liable to see the sort of demand surge for agricultural fiber that causes nonlinear behavior in coupled systems). Moreover, agricultural biomass has different characteristics than wood biomass (higher silica content, for example), and a shift might result in substantial increases in chemical waste emissions to the environment, at least while firms moved down the learning curve (that is, until incremental improvements to processes and technologies lead to environmentally efficient and mature "paper from ag waste" manufacturing techniques and practices).

   In many cases, of course, technological evolution is already occurring as a result of powerful cultural and economic forces, but it is still possible to try to evaluate potential implications for the environment and society so that the costs can be reduced and the benefits augmented. Thus, for example, it is apparent that e-commerce in its various guises is continuing to expand, and that each form of e-commerce,

such as B2B (business-to-business) or B2C (business-to-consumer), has its own set of social, environmental, and economic implications. Moreover, more recent developments, such as the shift from paper media distribution (e.g., newspapers and printed books) to electronic alternatives, tend to display shorter and shorter penetration times before they become widely accepted. Although it is daunting, and there are obvious limits imposed by the unpredictability and complexity of new technologies, it is still possible to do high-level studies of such systems before they are introduced. These can be viewed as scenario-generation processes (Box 10.3) rather than anything more definitive, with the purpose of providing more option spaces for responding to unanticipated emergent behaviors as the system develops. If such scenarios are to be useful, however, they should be developed before introduction of the technology being analyzed. Changes to technology become much more difficult because of, for instance, technology or social lock in, development of standards, network economics, and other phenomena once they go mainstream.

Technology lock in occurs when economic, cultural, and coupled technology systems coalesce around a particular way of doing something. An example is provided by the automotive industry, where a fuel cell propulsion technology is feasible today, but the energy supply infrastructure necessary to support it is not. The gasoline internal combustion engine is thus "locked in." Indeed, efforts to introduce plug-in vehicles are gated not by the vehicle technology itself, but by the need to introduce supporting infrastructures and the smart grid—in other words, to overcome the lock in of internal combustion technology. It would be difficult to introduce computer chips that are much more simpler than today's models, because the accompanying software products have been developed based on the capability of today's chips. Software and hardware both assume a particular capability in the other system, thereby locking each other in. As these examples illustrate, lock in does not imply that technological change is impossible, merely that it is much more difficult, and it may be guided by past decisions and technology adoptions.

Standards arise when technology systems must work with other technologies (in which case the standard governs interfaces between technologies), or when standards are necessary for the technology to link more widely. An example of the latter is rail gauges, where interconnectivity of rail systems both within and between different countries (or regulatory entities) required standard gauges (spaces between the two rails). Such standards may also be used strategically. The Finns, for example, in order to help protect their security, deliberately chose rail gauges different from the Soviet rail system with which they were interconnected, thus

inhibiting the potential value of the rail network for an invasion—a not unimportant choice, given how the Prussians used their railroads. Network economics arise when the value of a technology is disproportionately enhanced by the expansion of the network of users of that technology—the telephone or social networking services being classic examples. The more people on a social network site, for instance, the more useful it becomes for everyone.

3. It is characteristic of complex systems that the network that is relevant to a particular analysis is called forth by that analysis. Accordingly, *it is critical that the sustainable engineer be aware of the particular boundaries within which he or she is working, and be alert to the possibility of logical failure when one's analysis goes beyond the boundaries.* For example, to do systemic design or engineering work on New York City's water supply by considering only the five constituent boroughs of New York would result in a flawed assessment, because the system being analyzed (water provision to the city) is not adequately mapped by the political boundaries of the city. The relevant network and the boundary selected are discordant. Similarly, boundaries may be chosen on the basis of ideologies, disciplinary habit, or simplicity, rather than reflecting real-world couplings and linkages through times, again resulting in flawed perception and action. The bottom line for a sustainable engineer is that you have a responsibility to know what fundamental assumptions underlie your activities, especially in ESEM situations, and that choosing them and the boundaries they imply is an art, not yet a science. Moreover, many people have incentives to violate this principle. Consider, for instance, the idea that a climate scientist should manage information on climate change that reaches the public in order to reduce "moral hazard," that is, the possibility the public will continue to display consumption patterns that the scientist considers inappropriate. Assuming the scientist is competent, one would accept the scientific conclusions of such an individual, but should that authority extend into the domain of ethics, personal behavior, and lifestyle choices, and the ability to select them for others without their consent? It is not that the scientist should feel unable to comment on such issues, of course, but if commentary becomes manipulation of information availability or, more subtly, research funding, questions might arise. The boundary of scientific understanding does not extend to include political authority to make cultural choices.
4. This leads directly into the next principle. *There is a difference between social engineering and technical engineering*, and the sustainable engineer should not only understand, but should respect, that important difference. This is difficult terrain for sustainable engineers, who tend to be educated in quantitative methodologies, and for whom issues of social

engineering are usually opaque at best, so a few examples might be useful. For instance, some note that human rights standards—a form of social engineering—tend to be intuitive to many people who grow up in Europe and the United States, and to the Western NGOs that lobby for them. In some other cultures, however, they are sometimes seen as a form of cultural imperialism (an effort to impose one's cultural perspective on another culture that may not share the same values and perspectives). This is not because they oppose human rights per se, but because they believe the individualistic form demanded by Western NGOs is corrosive of other values. Thus, (then) Malaysian Prime Minister Mahathir bin Mohamed in 1997 marked the fiftieth anniversary of the Declaration of Human Rights by urging that it be revised or, better yet, repealed, because it focused too much on individual rights and not enough on the rights of society and the common good. Alternatively, suppose that one wants to design a consumption system that results in less emissions of carbon dioxide to the atmosphere, which could certainly be taken as a form of sustainable engineering. One could design it directly by, say, creating hybrid automobiles, banning all traditional internal combustion automobiles and telling people why it is being done (to reduce global climate change forcing). Such efforts would probably fail, at least in the United States. That such direct approaches can be politically unpopular explains why "hidden" social engineering is so common. Some argue that the global climate change community creates an air of crisis deliberately not because they believe dramatic change is an immediate threat, but because they want to change consumption patterns they regard as unsustainable and unethical.

In politics, this is actually fairly common. European governments blame European Union decisions for forcing reforms that they know they must make anyway, or U.S. legislation establishes goals that are unreachable but popular (such as requiring that all surface waters be fishable and swimable—who can oppose that?). Sustainable engineers, however, should bear in mind several real-world constraints on social engineering through technical engineering. Most important, perhaps, if the public or the affected group realizes they are being manipulated by engineering (or scientific) elites, it is highly likely that the effort will fail. Moreover, trying to socially engineer people on the sly runs the risk of undermining the credibility of engineering and science in the long run, which may in fact be happening in the climate change dialogs. This is particularly unfortunate because it is through engineering and science that the data and tools necessary for adaptation to changing conditions are developed; if credibility is lost, so is social resiliency. A high price indeed for trying to force people to live the way you think they should.

5. A critical understanding in the anthropogenic world is that *the sustainable engineers are also part of the system they are purporting to design, creating a reflexivity that makes the system highly unpredictable and, to some extent, perhaps more unstable.* This is generally obvious from the discussions in the previous chapters, but it has particular implications when the sustainable engineer is dealing with ESEM. It must be remembered that ESEM deals with systems not susceptible to easy fixes; earth systems are not projects to be completed, but conditions to be continually managed and designed, and that are always changing in response to the engineers' decisions. Climate change, for example, is not a "problem" to be "fixed," but a condition arising from a planet with 7 billion people on it—and it cannot be understood as a problem to be solved and forgotten. This bifurcates sustainable engineering. Most engineering still involves artifacts, which must be designed to be robust given their economic, social, cultural, and environmental context, but ESEM requires ongoing and highly sophisticated dialogs with the systems at issue. For example, a civil engineer may still design a levee in the Everglades and walk away, but an engineer responsible for the Everglades itself must maintain a constant dialog with that system. (The levee, for example, may have unforeseen effects on water flow in other parts of the Everglades, and thus need to be modified or enhanced over time.)
6. *Sustainable engineering at the earth systems level necessarily includes macroethical and worldview implications.* To achieve long-term clarity and stability of understanding and policy, such normative elements must be explicated and embraced, rather than being hidden. Moreover, it is likely that, in part reflecting the reflexivity of the system, values and ethical systems will at least in some cases be contingent, requiring that they be revisited as part of ESEM processes, instead of being fixed at some arbitrary point in time. What this means is that the sustainable engineer involved in an earth system—say, recreating the Iraqi Marshes, designing the machines to suck carbon dioxide out of the ambient atmosphere, or developing technology that grows meat from stem cells rather than on the hoof—should be sensitive to the fact that what she or he regards as fairly straightforward and unproblematic may be viewed by others quite differently. (Genetically modified foods from the United States, or Brazilian, as opposed to the European, perspective, is one obvious example. See Box 2.1.)

14.5 ESEM GOVERNANCE PRINCIPLES

The complexity that characterizes the Anthropocene has obvious implications for governance systems. The bastion of international governance has been the nation-state, at least since the Treaties of Westphalia in 1648 ending the Thirty

Years' War, but the dominance of the nation-state is being eroded by the evolution of large transnational firms and NGOs; global patterns of technological evolution and economic activity; the evolution of trans-state human rights which, at least potentially, are greater than (nation-state based) citizenship rights; the evolution of information networks and infrastructures which enable the creation of new communities at all scales; and, of course, the regional and global scale of physical, chemical, and biological systems impacted by human activity. That certain aspects of economic and cultural systems are increasingly globalizing, however, does not imply, as some believe, increasing homogeneity or simplicity: rather, the result is increased complexity at all scales. For example, the importance of the nation-state is not diminished by the evolution of trans-state institutions and patterns of communication, because nation-states establish identity at the level of personal and public moralities; represent and maintain cultural systems that, in turn, support individual freedom, economic development, technological capabilities, and the like; and maintain stability within their borders and national security beyond them. The role of the state is still critical in providing public goods, such as property rights, legal and personal security, honest government, and basic education, and states are still a major and necessary basis for the regulatory and institutional structures that support human and civil rights, and environmental protection. That is why some countries, such as European and North American states, remain consistently attractive to migrants. But nation-states are less and less credible as the sole, or in some cases even the major, source of sovereignty and authority in the modern world. The sustainable engineer who believes that established nation-state governments are her or his only stakeholders is in for a rude awakening.

Such complexity suggests that explicit international government mechanisms such as treaties function best as partial solutions to easily identifiable issues (that approach worked with global reductions in CFC production and use, for example: easily identified compounds, few producers, relatively few users, and easily available substitutes). But with more complex issues, sustainable engineering is harder, because there is no single source for reliable information on the relative importance of social and cultural values: Different nation-states display different values and priorities, and NGOs are far more active and varied than in years past. But this does not mean that the alternative is abdication of ethical and operational responsibility for outcomes. Rather, it is important to remember that "governance" is not "government": one is a dynamic process characterized by learning and evolution, the other is an institution. With this in mind, we can suggest a number of principals that, even if imperfect, can guide the sustainable engineer through this difficult maze.

We can begin by suggesting several overarching generalities. First, seeking certainty and complete solutions is a blind alley; human economic and

technological systems will continue to evolve and the status quo will continue to change. This should lead us to be skeptical of ideological solutions, although relevant ideologies are almost always powerful inputs into decisions. Second, we must never make the mistake of assuming the system isn't evolving unbeknown to us in any event. Thus, for example, every failure in the climate change negotiating process, from the Bush administration announcement that it would no longer support the Kyoto global climate change protocol to the failure of talks in Copenhagen in 2009, generated solemn statements by environmental leaders and climate change activists, duly repeated by the press, that efforts to respond to global climate change had failed. And certainly it was true that such events represented at least temporary slowdowns of organized state-to-state efforts. But of course adaptation continued, as it continues today, in many ways across numerous human and social systems. Private firms continue research into technologies, from telework to active carbon sequestration to genetically engineered crops, that reduce global climate change forcing. Environmental groups, if anything, accelerate their activism. Bilateral contacts between countries continue and are being strengthened. Investment in green power options continues, even in recessionary economies. New techniques, such as geoengineering, slowly became part of the climate change dialog. In short, a principal and visible activity ran into roadblocks, but to draw the conclusion that human adaptation to global climate change therefore stumbled was inaccurate. It arose only because the complexity of the human systems involved rendered them invisible: Those participating in the process were unable to perceive, much less understand, the underlying systems involved. The more generic conclusion to be drawn from this episode is that an anthropogenic world will be characterized by patterns of response and adaptation that are neither centralized nor evident—and that learning to perceive these subtle but very real patterns will be a significant challenge to sustainable engineering in itself.

An unprecedented and largely unperceived complexity is thus the background against which any governance models must be understood. Even at this point, however, that does not preclude development of some suggestive governance principles.

1. Perhaps the most important principle is that *conditions characterizing the anthropogenic Earth require democratic, transparent, and accountable governance.* As this discussion has made clear, the anthropogenic Earth and ESEM initiatives by definition raise important scientific, technical, economic, political, ethical, theological, and cultural issues in the context of an increasingly complex global polity. Given the need for consensus and long-term commitment, the only workable governance model is one that is democratic, transparent, and accountable. There are, of course, a

number of practical barriers to this principle, which make it more of an ideal than an implementable policy, at least at this point. Most obvious is that the principle suggests an administrative nightmare. The United Nations functions slowly and fitfully at best, yet it only includes countries as members. Imagine what an institution that included companies, NGOs and activists, online and geographical communities, and subgroups of countries would look like. There is also the important issue, which sustainable engineering inherently struggles with, of the different worldviews and languages that characterize different academic and professional disciplines. This is not merely unnecessary confusion. It is not the case that physics or engineering explains anthropology or sociology or literature, but it is also not the case that sociology or literary criticism engineers airplanes.

2. *Multiculturalism and dialog.* One of the things that is very difficult for the sustainable engineer to remember is that difference of opinion is healthy, because it enables the problem solver to determine what the real conflicts are in a particular system, be it a toaster or the Everglades, and to seek solutions that are truly stable because they speak to what different stakeholders regard as important. It is true that those with strong ideological predispositions will seek (to force) unanimity of opinion, but the sustainable engineer cannot take this route and remain true to the command to seek to solve real-world problems in workable and stable ways. Indeed, in some ways this is a significant strength of sustainable engineering, because periods characterized by a certain amount of conflict and tumult seem to lead to creativity and innovation, while, conversely, periods of orthodoxy are notably lacking in creativity. Indeed, such a description fits the Enlightenment well—and engineering, a supremely practical profession, exemplifies Enlightenment rationality. Unnecessary conflict may be wasteful, but a degree of conflict and difference of opinion can be catalytic (compare the Everglades example, with frank disagreements among the stakeholders, to the Aral Sea example, where unanimity was imposed by an authoritarian state).

   Sustainable engineering is in fact coevolving with a world that is increasingly diverse and multicultural, and that cannot simply be forced into one orthodoxy. Previous practices, such as the "high modernist" tradition of urban engineering and policy formulation that held until fairly recently in many cities, have given way to a complexity of voices and power that are a marked difference from the past. It is not just that the simplicity of high modernism would most likely fail in today's world because it would not lead to politically viable decisions. It is also that many of the powers of the anthropogenic world—the developed countries of the West, the up and coming countries of the BRIC category

(Brazil, Russia, India, China, and others such as Mexico and Indonesia)—are well aware of the power and importance of engineering and subsequent cultural capability in technology systems. The competition to best evolve technology, human institutions, and cultures is thus already heated, and such evolutions do not occur where stasis and orthodoxy have been successfully imposed. Multiculturalism is therefore not just a fact of life for the sustainable engineer, but a positive support for social and technological innovation and strength. Of course, when multiculturalism leads to social breakdown or chaos, as during periods of total war, or where unrestrained individualism reigns (similar to warlords in Somalia), it is counterproductive. Thus, there is a need for balance between chaos and rigidity, between individualism and communalism.

So, maintaining diversity of opinion and rigorous dialog, and a multicultural global polity, is not a luxury or an unfortunate necessity, nor is it simply political correctness; rather, it is a critical source of growth, evolution, and future socially and environmentally acceptable prosperity. Of course, there will continue to be many activists, ideologues, and fundamentalists of various stripes who will not be willing to accept dialog and disagreement as legitimate. As stakeholders, these must be part of the input that a sustainable engineer considers, but they should not be allowed to dominate design and engineering decisions, even if as an individual the engineer agrees with such positions.

3. *Engineering, management, policy, and governance are a continuing dialog, and sustainable engineers must understand themselves as part of the relevant system.* ESEM governance models, which deal with complex, unpredictable systems, must accept high levels of uncertainty and reflexivity as inherent in the subject area, and view ESEM, and policy development and deployment, as a dialog with the relevant systems rather than a definitive end point. ESEM governance structures should accordingly place a premium on flexibility and the ability to evolve in response to changes in system state and dynamics. The sustainable engineer, the earth systems manager, and the policy maker are part of a constantly evolving rather system, rather than standing outside the system guiding or defining it. This is, by the way, one of the problems of the "stewardship" idea: it considers the human to be apart from the system being managed, a "steward" of it, rather than an integral part of the system.

   This illustrates a major challenge for the sustainable engineer: the complexity inherent in working on physical and built systems whose time and space scales are much longer than the funding and policy scales that we are used to. In some cases—climate change or shifts in the nitrogen cycle, for example—science and policy mechanisms are simply too shortsighted and uninformed. Rather than respect the disruptive effects

of emerging technologies, smooth curves are drawn extending over a century or more that are interpreted as having predictive power. Such curves are simply wrong. On the other hand, with certain other issues, such as the legal implications of social networking technologies or the rise of virtual cyberspaces as separate economies, policy and legal institutions are simply unable to react in time; they are unable to couple to the much more rapid change that they purport to address.

The effective time cycles of engineering and policy formulation are also modulated by the perceptual and psychological horizons of individual humans and their governmental institutions. Those horizons are generally inadequate to the task of responding to the challenges of governance of an anthropogenic Earth. Figure 14.1 illustrates this issue with regard to environmental systems, where many of the perturbations of major concern tend to fall well beyond the temporal and demographic reach of traditional policy and legal mechanisms, and the human psychological horizon, which tends to extend to about six or seven years, and perhaps to the boundaries of the tribe, local culture, or nation-state (depending on the size of the country). Clearly, it is hard to maintain

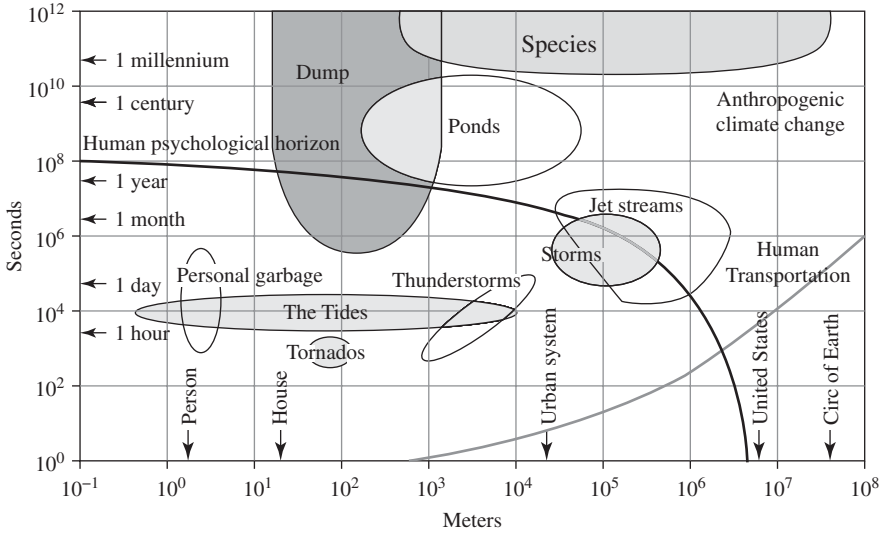


FIGURE 14.1 Human Psychology and Natural Systems Scale. This figure illustrates one of the difficulties inherent in earth systems engineering: The temporal and physical scales of the systems involved tend to extend beyond the psychological horizons of the individuals attempting to understand and manage them. This is not a barrier to rational and ethical earth systems engineering, but it is a factor that must be borne in mind, especially as people have a tendency to assume current conditions and technology will remain fairly stable into the indefinite future (as in the smooth curves characterizing many climate change predictions).

any sort of consistent sustainable engineering of such systems over periods of centuries, in the face of radical social, environmental, and technological change. The obvious example is the effort to find ways to manage long-lived nuclear waste. Where the environmental concern may last centuries or even millennia, and most countries are lucky to last a few centuries, one faces obvious practical challenges to continued institutional dialog.

There is an obvious governance gap when the cycle time of the perturbations of greatest concern—changes in patterns of biodiversity; changes in climatic, hydrologic, and oceanic systems; perturbations to fundamental elemental cycles; social and institutional adjustment to new technology systems—fall at the far end of the cycle time of even the fundamental human institutions of religions and cultures. Although there may be ways to manage such situations—for example, derive quantitative goals that link to desired end states, and develop policies to reach and maintain those goals—the global governance systems with which sustainable engineering must work are clearly not ready for such structures yet.

4. *Ensure continuous learning.* It follows from the complexity and unpredictability of the systems involved that continual learning at the personal, technical, and institutional levels must be built into the ESEM process. There already exists some experience with this approach. Many existing "high reliability organizations" such as aircraft carrier operations or well-run nuclear power plants usually have explicit learning structures. Adaptive management (see Chapter 13) in cases involving complex natural resource regimes such as the Baltic Sea, the Everglades, the North American Great Lakes, or particular forest or fishery resources, is heavily dependent on continual learning (and the relevant literature identifies cases where lack of this ability makes policy and social institutions quite brittle as change occurs). This learning process is messy and highly multidisciplinary, but it is particularly critical with ESEM projects, which in many cases will involve significant experimentation and a highly interactive relationship with the systems at issue.

14.6 ESEM DESIGN AND MANAGEMENT PRINCIPLES

There are a few important principles that don't fit well within either the theoretical or governance categories, which have to do with how one might interact with the integrated human/natural/built earth systems of the Anthropocene in the hopes of affecting their trajectories in desirable ways. This is a tricky area in part because complex adaptive systems can't really be designed as an airplane can, for example, because they will inevitably behave in ways that were not anticipated. Nonetheless, there are times when immediate action may be necessary to avoid disaster, or when the outcomes of inaction have a very high

cost—particularly when inaction results in irreversibility of undesirable characteristics. If you see a situation like that, but action is not being taken, it probably means that rhetoric is outrunning political will; this is something, if you are doing the sustainable engineering around a project, you need to be able to understand.

For instance, many developing countries, needing immediate augmentation of their energy systems, are currently installing coal-fired power plants. Technology exists to capture the carbon dioxide produced in these plants, liquefy it, and inject it into underground formations where it will not contribute to global climate change. In addition, new build carbon sequestration imposes bearable energy and economic costs; retrofitting tends to be much more expensive both financially and energetically. Under these circumstances, it would make sense if control of carbon emissions is a serious goal to at least plan and site the plants so that sequestration technology can be easily installed; instead, they are generally being built without any such considerations, which means that, given the half-century life of such facilities, they embed carbon emissions in the global economy for at least decades. If decision makers really believed that climate change was a serious and irreversible risk, they would be taking at least such minimal action as insurance; that it is not generally happening indicates that, despite the rhetoric, climate change is not yet viewed as a major problem. If you were asked to design a coal-fired power plant for such a country, and you designed it with (more expensive) carbon capture and sequestration capability, you would be making a mistake because, regardless of how you might prefer such an outcome, it's not what the client and the real-world conditions would be ready for. (As a sustainable engineer, however, you might encourage such foresight, show how it could be achieved, and at the least try to site and design the plant to make it easily upgradeable.)

So what are some ways to do sustainable engineering with such large and complicated systems?

1. An important point with unpredictable systems is to know from the beginning what the desired outcomes are, and *establish metrics that determine whether the system is indeed moving along an appropriate path to achieve the desired outcomes.* Concomitantly, potential problematic behaviors of the system can be anticipated and metrics established to determine if they are becoming more likely, or occurring, and if so at what level. For example, there is nothing wrong with corn-based ethanol as a technology; it is only when it is rapidly scaled up that it creates major social, economic, and operational dislocations. So one might have said, "Let's scale up corn-based ethanol until we find farmers shifting out of other crops and into growing corn at large scale, because that's

the point when economic and political factors begin to come into play." Such a metric, even though somewhat informal, might have avoided much of the turmoil that the U.S. policy on biofuels, focused on corn-based ethanol, actually caused. Such metrics are a way of institutionalizing a dialog with the system so that completely unanticipated behavior can be identified, and if it is high risk or has negative effects, mitigated before it becomes irreversible. Note, however, that some emergent behaviors, such as CFCs depleting stratospheric ozone, will be difficult to even identify a priori.

2. It is important to remember at all times that, unlike simple systems, *complex adaptive systems cannot be centrally or explicitly controlled.* It therefore becomes important to understand not just the substance of the system—the biology of the Everglades or the Baltic, for example, or the economics of corn-based ethanol, or the physics and chemistry of the troposphere—but also the inherent dynamics of the system structure itself. Each system will have some sort of diffuse control system, and if one wishes to shape its future path, one will need to understand where and how these control functions can be manipulated, and what the effects are likely to be. As part of this understanding process, one should try to identify the leverage points where small shifts propagate across the system as a whole, as opposed to other points in the system where any change tends to be dampened out by system structure. The famous example of the butterfly that flaps its wings and causes a storm elsewhere in the world has become an iconic illustration of chaos and complexity. What is perhaps forgotten is that millions of butterflies flap their wings thousands of times each day, without causing an ensuing storm. Perhaps the really interesting question is why one flap may have such an impact, when all the others don't.

3. *Premature lock in of system components should be avoided where possible.* In complex systems, artifacts, technologies, and choices among alternatives can get quickly locked in, depending on the structure of the particular system or network. Moreover, tightly coupled networks are more resistant to change than loosely coupled networks. These characteristics of complex systems explain, for example, why changes to pollution control equipment regulations are more easily accomplished than changes to product design or manufacturing process regulations. In the former case, the technology is only loosely coupled to underlying product and manufacturing networks, and thus can be changed with only minimal implications for other aspects of the product and manufacturing networks; lock in exists, but is not significant. In the latter case, however, the intervention is to tightly coupled networks (product design and manufacturing processes are complex, tightly coupled, interdependent networks), which

means that changes and shifts in state propagate throughout the system. Lock in of changes to such systems occurs more easily, and is more irreversible, in the sense that changing them once implemented is much more difficult.

4. These characteristics of complex systems impose a substantial challenge to the sustainable engineer, because it is an important principle that, *whenever possible, engineered changes should be incremental and reversible, rather than fundamental and irreversible.* Accordingly, lock in should be avoided (because it leads to irreversibility), in part by assuring that, where couplings to other networks do exist, they are designed to be as loose as possible, and as few as possible. After all, if the system begins to behave in an unanticipated and undesired way because of the action taken, easy reversibility is a very valuable characteristic. Moreover, as a fundamental technology system is implemented, scale up should anticipate nonlinear behavior on the part of the system; emergent characteristics should be assumed to be the rule, not the exception. This principle is particularly important where large engineering projects with potentially regional and global scale earth systems engineering implications are involved—as in, for example, many geoengineering projects. Similarly, one initial advantage of corn-based ethanol technology is that its massive scale up is reversible, at least until the relevant production and distribution infrastructure is built, and large firms and constituencies develop significant economic and political interest in keeping the technology at scale, at which point technological and political lock in begins to take over. Another way to approach this principle is to think about how an ESEM project, say, a large mine, can be conducted responsibly and when at end of life can be closed in such a way as to replicate initial starting conditions as closely as possible (or desirable: because the end-of-life landscape is a design project, it can include desired improvements to initial conditions as well; see Box 14.1). Thus, a well-designed mine is reversible in land use terms not in immediate timeframes, but over its lifecycle.
5. Earth systems engineering and management projects should aim for *resiliency, not just redundancy, in design.* Redundancy provides back-up capability in case a primary system fails, and is commonly designed in to high-reliability systems such as jet airplanes. Redundancy assumes, however, that the challenge to the system is of a known variety. Resiliency, to the contrary, is the ability of a system to resist degradation and, when it must, to degrade gracefully even under unanticipated conditions. Neither is easy to achieve. Designing redundancy requires enough knowledge of the system to identify (likely) failure modes. Designing

BOX 14.1 Mining as ESEM

The mining industry is of interest to sustainable engineers for several reasons. First, it expresses a serious internal conflict in the sustainability community (and hence in the input to engineering problem solving). Some sustainability advocates, for example, hold that no mining, by definition, can ever be sustainable, and that any process that takes materials from the Earth is by definition unsustainable. Others, observing that current technology and economic systems are dependent on mined minerals and materials, hold that mining will continue to be done, so it should be done in as socially and environmentally efficient a manner as possible. Although the former position has the advantage of being ideologically pure, it is infeasible for sustainable engineers who need to create real-world solutions and thus accept that mining is a part of the current paradigm.

More subtly, mining involves a cultural conflict as well. Mining firms often operate in dangerous and remote parts of the globe, and mining companies thus tend towards institutional cultures that are rough and ready, and not easily coupled to the concerns of many NGOs. This inherent cultural conflict is complicated by a strong difference in approach between large transnational mining firms, which tend to pay more attention to global human rights and environmental concerns, and smaller, more informal firms, which in many cases coalesce around a particular ore body or opportunity, exploit it as cheaply as possible, and then terminate operations, often leaving significant community and environmental costs behind them. Moreover, in many poorer areas individual artesian miners try to exploit the tailings and any minor opportunities to generate income; such mining is often highly hazardous, usually corrupt, and not infrequently associated with long-term conflict, especially in Africa.

Thus, over the past decade or so there has been a concerted effort by big transnational mining companies to reach agreement on how large-scale mining activities can be undertaken in a way that is socially and environmentally responsible. For example, the World Business Council on Sustainable Development sponsored a large project, the Mining, Minerals, and Sustainable Development Project, that issued a substantial report in 2002 titled "Breaking New Ground." Some mining companies now issue sustainability reports, and mining remains a very intense focus of many NGOs, both those that deal with human rights and those that deal with environmental issues. Obviously, the companies and their NGO critics do not always agree, but it should be remembered that it is only the first-tier firms that are engaged, and that in many places it is the small and family firms, and the individual miners, that give rise to the more serious concerns, especially on a per unit produced basis.

A large mine operation is in many ways a small ESEM project in itself. Consider, for example, the Grasberg Mine in Indonesia. One of the world's largest producers of copper and gold, the mine operation extends some 100 miles (approximately 160 kilometers) from mountaintop mine to the mine's port, linked by tram, road, and helicopter service; the facilities extend some 50 kilometers across in some

areas. Total land involved in potential mining operations extends over 2.24 million acres. The mine not only operates its own processing and milling facilities, but energy production capabilities as well: it imports 700,000 tons of coal per year to feed its 195 megawatt coal fired power plant, in addition to 100 million gallons of diesel per year for equipment use. The facilities also include a lime plant, hospitals, stores, schools, and towns and camps—the town of Tamika, for example, didn't exist until 1970 when the mine began operations and now has a population of over 150,000. In addition, the taxes paid by the mine provide some 2.5 percent of the GDP of Indonesia, and 49 percent of the gross domestic product of Papua, where the mine is located. Security is provided by the Indonesian military and is somewhat complicated, in part because different elements of the Indonesian government seek to control various aspects of the mine's revenue streams. In terms of production, over the life of the mine some 3 billion metric tons of ore will be recovered, requiring the movement of some 3.3 billion tons of overburden. Regardless of whether one approves of mining, or of this particular mine, it is clear that in terms of local and regional cultural and economic impact, of physical activity, of quasipermanent change to local systems, and of contribution to world material flows and consumption, such a mine is clearly a major factor. It is, in other words, ESEM.

The Grasburg Mine is a very large project, but virtually all significant mines regardless of where they are have substantial effects on local and regional communities, environments, and infrastructures. (All mines require substantial infrastructure support, especially in terms of energy, water, and transportation of various kinds.) Thus, designing and operating such projects is always an interesting challenge to the sustainable engineer. Moreover, mines by their very nature are not inherently "sustainable"; the ore body is, after all, limited and at some point, no matter how large it is, it runs out and the attention must shift to reclamation of the area for other uses. This offers another opportunity for sustainable engineering: rather than simply closing a mine and remediating its potential environmental impacts, one might want to design it from the beginning to shift to other uses. For example, mines usually are fed by high-power lines because of their demand for power; this means that if geothermal or large-scale solar could be installed in the mine location upon closure, it could (a) continue to offer the local community employment and taxes, (b) take advantage of preexisting infrastructure, and (c) place power generation facilities in areas that are already disturbed. In fact, as attention to sustainability increases on the part of society and NGO groups, mining and other extractive sectors could turn out to be important areas of practice for sustainable engineering.

resiliency also requires a high degree of knowledge, in part because resiliency is not a global characteristic of a complex system, but is generally defined in terms of a capability to resist a particular class of insults or perturbations. Thus, it is important when working with a system—say,

trying to help an urban region improve its ability to resist and respond to terrorism, while at the same time creating better quality of life for its inhabitants—to define resiliency in appropriate terms, so policies can be evaluated and resource allocations made, with some idea of associated costs and benefits.

14.7 SUSTAINABLE ENGINEERING, ESEM, AND THE FINAL PRINCIPLE

These ESEM principles are generic and incomplete, and will certainly be augmented and improved going forward. Moreover, they are both fairly general, and in some cases difficult to achieve in practice. Making sure all affected groups have a voice in decision making, for example, is difficult even for small sustainable engineering projects, but infeasible for many regional and global ESEM activities. In the case of climate change, for example, the negotiation process is already quite unwieldy, yet arguably a major drawback of that process is its domination by only a few worldviews (e.g., environmental).

Against these quite legitimate challenges to various principles is the simple fact that sustainable engineering and ESEM are not important today because of whim or fantasy, but because they are the only path towards rational, responsible, and ethical engineering and management in the complex world that humans have already created for themselves. In this light, it is a hopeful sign that there is so much existing knowledge and wisdom that can be integrated so smoothly into sustainable engineering even at the largest scale, that of the planet and its global economic, technological, social, cultural, institutional, and natural systems.

This leads to the final point: In a complex regional or global earth system that is both unpredictable and rapidly changing, the best adaptations are to embrace the development of option spaces and to be prepared to muddle through, rather than rely on any explicit and formal path. For many engineers, a constantly changing, somewhat arbitrary process that responds to changes throughout the system in different ways—some social, some institutional, and some technological, for example (think climate change adaptation)—seems messy and suboptimal but, where the systems are as difficult and contingent as they are today, and so interconnected with each other, and when we know so little, well, inspired muddle is often the best path. Helping that along is the deliberate effort to create options, so that one has a palette from which to choose.

One example of this strategy is the process by which the electronics industry reduced its use of CFCs in response to the Montreal Protocol, which banned such substances because they reduced the stratospheric ozone

layer—important because it absorbs high-energy ultraviolet light before it can affect biological systems on the Earth's surface. Used mainly for cleaning purposes, CFCs were very effective and almost universally relied on. Over the years, however, a number of alternatives were developed, ranging from cleaning systems using other chemicals such as terpenes, to water-based detergent systems. Accordingly, it was very easy technologically to switch to other materials without major production or quality disruptions. In fact, the biggest barrier to switching turned out to be the requirements in military specification documents that referred specifically to CFCs, because in order to maintain contract terms, each individual use—and there were many tens of thousands of such uses—had to be tested and approved. Not only does this example illustrate the benefits of a robust technological option space, but it also indicates the need when evaluating sustainable engineering alternatives in the real world to pay attention not just to economic and technological factors, but to institutional factors (e.g., military contract requirements) as well.

So what is the final principle, good for both ESEM and sustainable engineering? When dealing with highly complex systems, embrace rigorous and principled muddle, rather than seeking false, and ultimately dysfunctional, simplicity.

Exercises

14.1 You are asked to design a large bioenergy plantation in East Africa, out of land that is currently occupied by a few Masai tribes with a small number of urbanized settlements. You have a mandate to follow principles of sustainability in your design. What are the steps you would take to ensure that ecosystem integrity is ensured, that livelihoods are not endangered, and that the Masai, who tend to be nomadic, are not fenced in or out of the plantation area?

14.2 You are the United Nations sustainable engineering expert in charge of geoengineering policy. For purposes of simplicity, political expediency, and economics, you have been asked to select one geoengineering technology for global implementation. Your choices are: (a) lifting a large number of reflective aluminum balloons into the stratosphere to reduce incoming radiation; (b) putting a lot of iron into oceans around the world so phytoplankton will absorb all the excess carbon; and (c) immediately cutting fossil fuel use around the world by 50% in every country, no exceptions.

a. Applying relevant ESEM principles, which option would you choose, and why?

b. What are the drawbacks of each choice, and are they mainly political, social, economic, technological, or environmental?

c. Which option would you expect the developing country mem-

bers of the United Nations to support, and why?

14.3 A significant amount of iron ore is discovered in a tropical forest with vast reserves of exotic floral species that are expected to be of immense though as yet unknown pharmaceutical value. Several indigenous communities are becoming quite alarmed at the prospect of a multinational mining company making secret deals with the government and getting mining leases without having to make amends for the environmental and social losses that may ensue. You are writing a feature piece on this for *Nature* magazine. How would you communicate the central complexities of this case?

14.4 You are approached by a large mining company that has just discovered a major ore body in the western United States. As a sustainable engineer, you are asked to design a socially and environmentally responsible mining operation.

a. What would you identify as the critical lifecycle stages of the mine operation?

b. What issues do you see as critical in terms of integrating the mine into local and regional infrastructure systems?

c. Mine closure is always a difficult task. What recommendations would you make to reduce the environmental and community disruption that is usually entailed in this process?

Annotated Bibliography

1. There are as yet no books that I am aware of on earth systems engineering and management; the one that comes closest is Braden Allenby, 2005, *Reconstructing Earth* (Washington, DC: Island Press). A more recent article by Allenby, "Earth Systems Engineering and Management: A Manifesto," *Environmental Science & Technology* 41 no. 23 (December 1, 2007): 7960–7965, provides perhaps the most up-to-date summary of the state of ESEM. In terms of natural systems, there are more options, including V. Smil, 1997, *Cycles of Life: Civilization and the Biosphere* (New York: Scientific American Library). Vaclav Smil is a prolific and easy-to-read writer, and this book is a good overview of natural systems and their relation to human systems. It is neither highly technical nor so watered down that it fails to interest and educate, so it is a good resource.
2. Less accessible but still useful are the *Scientific American* special issue, 1989, Managing Planet Earth, *Scientific American* 261 no. 3; and the *Science* special report, 1997, "Human-Dominated Ecosystems," *Science* 277: 485–525. Both of these issues, as one would expect from these journals, are excellent. The *Scientific American* covers a broad range of issues, from forestry to climate, and remains a valuable read, although it is getting somewhat dated. It has the distinction, moreover, of containing the article by Robert Frosch and Nicholas Gallopoulos, "Strategies for Manufacturing," that many believe

was instrumental in beginning the study of industrial ecology. The *Science* special report is more technical, and more focused on biological systems, but its message—that "there are no places left on earth that don't fall under humanity's shadow"—is perhaps all the more powerful for that. These sources are useful both in themselves and as representative of how concern about larger earth systems, and thus ESEM, has grown from more traditional concerns about environmental and energy issues.

3. Regarding urban systems, among the most interesting books on high modernism remains Le Corbusier's classic, first written in 1929 and republished in 1987, *The City of To-morrow and Its Planning* (translated from the 8th French edition by F. Etchells, reprinted by Mineola, NY: Dover Publications). The two most interesting critiques of high modernism are Jane Jacobs's excellent [1961]1992 book, *The Death and Life of Great American Cities* (New York: Random House; reprint, New York: Vintage Books); and Charles Jencks's 1977 book, *The Language of Post-Modern Architecture* (New York: Rizzoli International Publishers). Peter Hall's 1998 book, *Cities in Civilization* (London: Weidenfeld & Nicolson), is long, but a wonderful romp through the urban core of the human species, while Berman's 1982 book, *All That is Solid Melts into Air* (New York: Simon & Schuster), is a good overall critique of high modernism generally.
4. A good introduction to geoengineering is the Royal Society report, *Geoengineering the Climate: Science, Governance and Uncertainty,* published in 2009. Much more broadly, Martyn Fogg's 1960 book, reissued in 1995, called *Terraforming: Engineering Planetary Environments,* is still a good introduction to some of the basic principles of terraforming. Martin Beech's *Terraforming: The Creating of Habitable Worlds*, published in 2009, is a nontechnical and very accessible treatment of the subject. Finally, there is a growing literature on not just space habitat engineering, but space habitat architecture; an example for those who are interested is A. S. Howe and B. Sherwood's 2009 text, *Out of This World: The New Field of Space Architecture (Reston: VA: American Institute of Aeronautics and Astronautics)*. Enjoy.

CHAPTER

15 The Engineer as Leader

You who choose to lead must follow,
But if you fall, you fall alone.
If you should stand, then who's to guide you?

—"Ripple," The Grateful Dead, 1971[1]

15.1 INTRODUCTION

The purpose of this final chapter is simple. It is to provide a few closing comments on how technocrats, including engineers of all stripes, should increasingly assume leadership positions in their society and to suggest some ways that they might perhaps approach that role. It will not, however, "teach leadership." Although there are many textbooks that purport to do so, serious leadership capabilities develop over an entire career and are a product of many things. In some ways, this entire textbook is, in fact, an effort to provide the initial grounding on which technologically aware and competent leadership skills can be built. Accordingly, this will be a short summing up, rather than some

[1]Lyrics to "Ripple" by Robert Hunter, copyright by Ice Nine Publishing Company. Used with permission.

sort of populist leadership seminar ("Now! For a Limited Time Only! The 27—Count Them—27—Steps to Becoming Rich, and a Leader, and Famous!"). For one thing, "leader" is a really fuzzy concept: the middle manager who guides his or her unit through a difficult period, the engineer who uses her or his spare time to work with homeless kids, the brave soul who gets elected to local office, the owner of a small consulting firm—all are leaders. The basic argument of this chapter, however, is that in a world increasingly defined by technology, society requires leaders at all levels who understand and are comfortable with technology.

Certainly, it has always been the case that some engineers have been leaders: George Washington was a surveyor; Thomas Jefferson was a designer, inventor, and architect; and Herbert Hoover was an engineer (by far the dominant profession of American presidents, however, is law). China's leadership today is dominated by engineers. Some firms are known for being led by engineers, as opposed to lawyers, accountants, and MBAs. But in general it is the case that engineers have functioned in most cases as technical members of institutions and teams led by other people with different professional training—business, law, or military, for example. This pattern reflects a number of factors, many of which arise from the way we teach engineering. As an undergraduate professional degree that involves enough necessary technical material so that the ancillary courses that would round out a potential leader—courses that would lead to development of good writing and verbal communication skills, some knowledge of culture and belief systems, an in-depth knowledge of economics and capitalism and industry, some sensitivity to institutional structure and dynamics, and perhaps most subtly a faith in one's ability to lead in difficult times and places—have no room in an engineering curriculum. This is a recognized problem, and it is changing somewhat (that's why you might be reading this book), but it is changing only incrementally and slowly.

This perhaps leads to the most important point about leadership, and we might as well get it over with at the beginning: Don't depend on your engineering education, or on anyone else, to make you a leader. That doesn't work. The institutions that will teach you engineering are good, but they are old, and they are increasingly obsolete in this rapidly changing, technologically complex, information dense world. Institutions and mentors can certainly help you, but they can only point the way and provide advice and wisdom. You have an individual responsibility here to develop and grow yourself. And a society that is increasingly characterized by complicated and rapidly evolving technologies and engineered systems badly needs you to accept that responsibility, because the lack of technically competent leadership is becoming more and more apparent, and more and more dangerous.

Bluntly put, being a leader is an inseparable part of sustainable engineering. So the first step is to get rid of that idea of yourself as just an engineer, or

just a technically trained expert, and to begin understanding yourself as an engineer and a leader in training.

15.2 ATTITUDE AND PREPARATION

Let's begin with an obvious but difficult point. The real challenge of sustainable engineering is the design, operation, maintenance, and evolution of engineered systems in a context of serious complexity, where the usual guides—simple design tools, assumptions about what society wants—are inadequate. This is especially true where different stakeholders have very different, perhaps even mutually exclusive, worldviews. Moreover, you are a human being yourself, so you have your own set of values and that, because your job is to create ethical and stable solutions, not to impose your own views on others that may disagree, you have learned to suspend as you listen to others, determine their concerns and agendas, and integrate them into the process.

This creates a difficult balancing act. As a technically competent individual, you need to respect the complexity and unpredictability of the systems you are responsible for or with which you work. But as a leader, part of your job is to understand this complexity well enough to be able to simplify it for others who may not have your background. Moreover, you need to do this with integrity, so that stakeholders can make decisions that are truly their own, not that are predetermined by you as the leader. This requirement for integrity is violated, for example, if you manipulate funding, or research directions, or data presentation, in such a way as to drive those with less knowledge in the direction that you want rather than in the direction that they may choose. Yet at the same time, you presumably have a vision based on your personal experience and values that motivates you and that you as a leader express to others in a way that motivates them as well. Stated this way, it sounds as if there is no solution, but in real life, leaders provide solutions all the time. You simplify, but do so with integrity; you express your goals and solutions, but do so in ways that show you have listened and learned from others and that you respect even those with whom you disagree.

Good leadership, in fact, requires continual learning. This does not just mean doing well in traditional engineering courses; actually, to the extent such courses imply that answers are generally found in the back of the book or in extremis, a result of identifying the right algorithm, they may actually teach you how not to learn in a meaningful sense. You can't afford that, nor can society. You can't afford it because one of the major means of trying to future proof yourself in an unpredictable and rapidly changing world is being able to ride change effectively. And that, emphatically, is not in the back of the book in the answers section. Society can't afford it because in the age of the anthropogenic

Earth, if the engineers and technocrats who are building the future are blind to its implications, the results are not just locally catastrophic, but could be devastating globally.

So you should practice learning at two levels at all times: the explicit learning level and the metalearning level. At the first level, you learn the techniques, thought processes, algorithms, and other tools and attitudes that characterize engineering—you *do* need to be technically competent. That has to be your baseline. But at the metalevel, you should constantly be questioning the class, the subject matter, and the institution: What is it *not* teaching you? What skills are you *not* learning? What is the professor saying that is already obsolete? What are you being taught that helps future proof you, and in doing so gives you the foundation to grow into leadership? What would you rate as the strengths and weaknesses of those around you who are considered leaders: the professors, the deans, the managers, the gurus? How would you rate the strengths and weaknesses of this book?

To provide a personal example of how this process works: One of the most difficult jobs I ever had was working as a lawyer for a legal vice president who considered anyone who wasn't at the office twelve hours a day six days a week a slacker. This individual was also, as one might guess, a micromanager and a difficult personality. And yet I learned more from that person than from any other manager I ever worked for. Part of it was that, despite the individual's shortcomings, he was very adept at office politics; I was extremely naïve at the time, so it was a revelation to me to watch him in action. In addition, critiquing his management style taught me an enormous amount about leadership—mostly what not to do, of course. For example, because the time expectations were unrelated to work productivity, several of my friends spent most of their day reading *The Wall Street Journal* and managing their investment portfolios (this was before the Net offered such varied distractions). Lesson: evaluate people on what they produce, not metrics such as "time at desk" that don't correlate with economic value. I watched brilliant but self-directed people wilt under micromanagement: That lesson was to understand your people and give them what they need to succeed in their job, not what fits your own psychology. Some people like micromanagement; it tells them they are on the right path and the attention encourages them to understand the importance of what they are doing. Others hate it; they want to be given a task, a time in which to complete it, and then be left alone unless they need "covering fire." ("Covering fire" is giving the people who work for you the necessary support and protection from office politics to get their job done; even in the best organizations, individuals will often run into situations that must be resolved at higher levels of management.)

And I saw what this individual gave up to maintain his position. If you reach a certain level, the organization expects you to understand that it has

first call on your time. If you have to come in on weekends and cancel vacations and end relationships, so be it. In the case at hand, I watched this individual do all of that. He had chosen to be a vice president, and that was the tradeoff. Lesson: Make sure you know what leadership is going to cost and that you understand the responsibilities as well as the benefits.

An important part of leadership, in fact, consists of knowing yourself and what you want from your work and your life. If you don't want to make the sacrifices of an executive position, learn that about yourself, and direct your energy and interests elsewhere. If you like working with nonprofits, or being part of your town's or state's government, perhaps as an elected official, or perhaps serving on advisory boards—do that. Part of getting an education is gaining the ability to create a life for yourself that is personally and socially worthwhile, and how you choose to do that is an important part of developing the self-knowledge that helps you become a leader in whatever area you do choose.

Or take the institutional level. Most institutions have public faces that consist of vision statements, the way they position themselves in advertisements or policies, the way they respond to public inquiries and concerns, and the like. They also have private agendas, which may or may not align with their public images. Enron, the famous failed energy firm, for example, had a public face as one of the strongest social responsibility programs in industry in public—and ran some very touching television advertisements bragging about their program—but in private rewarded very different behavior that often came very close to illegality. Understanding what really motivates institutions and what stage they are at in their evolution, is a critical part of leadership that seeks to introduce meaningful change.

For instance, suppose you are deeply committed to a more egalitarian, more environmentally responsible world. You are hired by a manufacturing firm, and you want to change their behavior to fit more with what you think the world needs. The first step is not to draw up a list of deficiencies and show them to your boss; that will usually not be regarded positively. Rather, first you learn, again at both levels. Learn what you need to know to do your job effectively and well—this gives you credibility. But also do your metalearning: What are the real values of the company? What language does the institution respect most: Legal? Financial? Engineering? Science? Social values? The answer will differ depending on whether the institution is an IBM, or a Body Shop or a Ben and Jerry's Ice Cream, or a national weapons laboratory. And then identify projects and opportunities that lie at the intersection of your interests in sustainability and the institution's interests, and figure out how to express them in terms of the language that the institution respects. In almost any situation, you can find ways to advance social and environmental goals, but you need to respect the institution you work for, and its values, to determine how best to do that.

This raises another important question you should seek to understand with any institution: How ready is it for fundamental change? If the institution is currently successful and is confident in its capabilities in the immediate environment, suggesting dramatic change is likely to be difficult and you may want to experiment with incremental improvements. Consider, for example, the manufacturing firm mentioned above. If it has just launched a successful line of products, it is highly unlikely to want to invest in a redesign in the short-term—when mainframes were selling well, IBM was quite happy with that product line, for example. Under such circumstances, you might want to suggest incremental improvements: Are there cleaning systems that use less toxic materials, for instance, but don't require product redesign? Can the packaging of incoming components and outgoing product be simplified? Are there ways to save money and increase productivity for the firm, such as virtual office arrangements, that also benefit the employees and the environment?

Remember an important principle of institutional fitness: An organization that is extremely competent at one point in time usually does so by making itself extremely competitive in a particular niche. This means that when conditions shift, it is likely to be caught flatfooted. A firm that puts most of its resources into being well adapted to one set of conditions will generally be much less fit when those conditions shift rapidly. And, of course, that is exactly the environment we are in today. So, if the organization is very successful, do incremental changes, but be prepared with significant changes in your pocket, because it is likely that it will need them. IBM was on top of the world when the world loved mainframes; when that changed, it took some fancy footwork to convert the firm from a primarily manufacturing entity to a primarily service entity. Similarly, when telephone technology dominated voice communications, AT&T was the world-class expert; when Internet Protocol technology became the underlying platform for voice, AT&T ran into serious financial trouble and ended up being bought by Southern Bell. The classic case is RCA—an extremely successful firm when vacuum tube technology dominated, but a failure when chip technology came in. Many traditional camera and film companies, including Polaroid, Kodak, and Fugi, are struggling to make the shift from film-based photography to digital photography. In cases like these, management practices were relatively fixed and hard to change when the firms were successful; but as they came under stress, new practices, such as teleworking and virtual office arrangement, could be introduced, or new products developed (e.g., Kodak digital cameras). So if a technocratic leader wanted to introduce changes to processes, products, or services—many of which tend to improve environmental performance, make employees and their families happier, and improve the productivity and financial performance of the firm simply because they're more efficient if nothing else—the time to do it was not when the firms were successful, but, rather, when they

were under stress. Know your institution. Know how to lead under different circumstances. And learn to judge what phase your particular institution is in.

And, of course, continue your education formally if you have a chance. Most engineers are relatively quickly promoted into positions of authority, such as project or team managers, or midlevel management positions in firms. These involve skill sets and handling challenges that are quite different than those that an engineering education of any type can help you with. (Do you really know enough about financial and accounting regulations to not unintentionally violate them? What about environmental law?) At some point, then, options such as an MBA or project management certificate may well become an advisable part of your portfolio. But the challenge is more profound than that, of course. The accelerating evolution of emerging technologies, and the social, cultural, and institutional changes that are usually entailed, mean that technocratic professionals face far more rapid and unpredictable change over their working career than at any previous time in history. Continuous learning, formal and informal, isn't a luxury in such an environment; it is, in fact, a prerequisite for continued relevance. It is not just important if you want to be a leader; it is critical if you want to keep being employable.

15.3 LEAD BY FOLLOWING

By definition, leadership often involves creating or supporting change, and most institutions and individuals don't like to change. This means that good leadership requires an ability to present change in the way most conducive to acceptance, or in many cases, to present change as if it weren't change. To take another example from industry: AT&T created a team in the 1970s whose job was to implement industrial ecology in the firm's manufacturing operations. The environmental experts on the team were heavily in favor of requiring all engineers to attend classes on environmental issues, and to be required to do environmental reviews for all products, using new methods and processes. This reflected their institutional function, of course—they were the repository of environmental competence within the firm and thus viewed it as the first priority. Trying to do so, however, would have instantly doomed the effort: Bell Laboratory design teams were already fully committed, and placing such onerous requirements on them, including new design processes, gates, and metrics, would have failed completely. Alternatively, the choice was made to just create a new design module for the existing design process (hence the Design for Environment module for the "Design for X" process). This had the same effect of injecting environmental considerations into the design process, but did so in a way that made it appear as simply an incremental extension of existing practice. Moreover, further investigation revealed a number of other levers that could be used to achieve environmental excellence without affecting the design

process. For example, most design shops use ordering lists of standard parts and materials, and integrate different design specifications and standards into product design. Changing these underlying processes and documents to make them more environmentally friendly had the effect of improving product design from an environmental perspective, but without requiring any action on the part of the design team. So packaging that required the use of CFCs, components containing polychlorinated biphenyls, mercury switches, and other materials and assemblies that had less environmentally problematic substitutes were removed from the standard component ordering lists. Specifications and standards were similarly reviewed and modified. The result was environmentally preferable design without any additional effort or attention required from the already stressed design teams.

Obviously, there will be times when the changes required are profound and cannot be accomplished except by direct and obvious means. If, for example, a firm has suffered a severe challenge and needs to significantly change its operations or focus to survive, change will be difficult and foundational no matter how one tries to sugarcoat it. Consider the history of Nokia, a major producer of cell phones and related electronics, and currently a mainstay of the Finnish economy. The firm began in 1865 as a pulp and paper company; it grew by, among other things, producing shoes and boots, and even cable for the Soviet Union after World War II. Only in 1960 did the firm, which was at the time a somewhat confused conglomerate, begin to migrate to electronics, and it wasn't until the 1990s that it divested itself fully of its nonelectronics businesses. Each of these shifts represented deep changes to existing and comfortable patterns, and each was made possible only by strong leadership at many levels of the firm. IBM's transition from a primarily manufacturing firm to a primarily service firm is a similar story. Leadership under such circumstances is difficult, especially where the future shape of the core technology is uncertain, and is usually best done as quickly and as deeply as possible. This is because a series of constant small changes, with the employees and the institution unsure of what and where they should be focusing, can be extremely debilitating.

Nonetheless, the general rule remains the same: Change only as much as you need to, and to the extent possible make the change incremental and, indeed, even invisible.

15.4 PERSONAL CHARACTERISTICS

Everyone's leadership style is different, but there are a few common threads worth mentioning. First and perhaps most important, develop the basic skills required by leadership:

1. Learn and practice good communication skills. Look for opportunities for public speaking and practice your writing skills, especially for nontechnical audiences. Write an opinion piece for your local paper, for

example. If you don't fully understand what you are trying to communicate, it's a sure bet your audience won't.

2. Learn and practice good organizational skills. Most important, be competent at what the organization wants you to do, whether it is an engineering task, project management, or customer interactions. If you don't have the experience or skills that you think you need, try to get them. Most companies appreciate your wanting to be well trained and competent, and will help you get what you need (indeed, many have formal education and training programs—use them).
3. Speak the language of the tribe you want to talk to, not your own. If you are speaking to financial folks, talk finance; if you are speaking to line managers, speak in terms of their responsibilities and interests. They, like you, might want to save the world, but they, like you, have an institutional responsibility that they must respect. Moreover, the more you use their language to communicate what you want, the more any changes you want to encourage appear incremental, familiar, and less threatening than might otherwise be the case.
4. Learn and practice leadership. If you have an opportunity to lead a team, or to manage a project, don't just focus on the project itself, but also pay attention to your performance. What did you do that worked well, and what didn't work, and why? Do you need extra training to be able to lead in that situation (for example, do you need a project management course)? Ask the people who work for you what they think about your performance. If they're afraid to tell you, that in itself tells you you're not doing well as a leader. If they do tell you, be grateful; if it's good criticism, be even more grateful. You don't learn from syncophants.
5. Don't sell yourself short. Many people fail, or don't live up to their potential, because they undercut themselves, either indirectly, or by never taking up a challenge for fear they may fail. Of course you may fail, but . . .
6. Don't be afraid to fail. Life is a marathon, not a sprint, and so long as you're careful to learn from your failures as well as your successes, you can continue to build your leadership potential even through—or sometimes especially through—failure. Be sensitive to your culture, however. In the United States, failure is not generally heavily penalized; in many less entrepreneurial countries, failure can be held against one.
7. Don't sell those who work for you short. The bumper sticker may read, "Success is ours; failure is mine," but there's truth in it nonetheless. Your firm will recognize a team that is productive and effective, and the more you make sure your team members are recognized and rise, the more you will be equally valued.
8. Listen to your critics and, indeed, seek them out; you will learn a lot more from people who critique you than people who agree with you. They may not be right, of course, and one of the skills you should work

on is knowing when to change your views and positions based on criticism, and when not to. Never changing in response to good input is not a sign of virtue; it is a sign of stupidity.

9. Learn when to negotiate and when not to. If someone raises legitimate concerns that haven't been considered, or is seriously interested in agreement and mutual progress, see what can be done. On the other hand, many times you will be faced by stakeholders who, for a number of reasons, either are not interested in real solutions, or cannot afford to be seen as giving any ground. Negotiation under such conditions will be ineffective and you will need to manage them by other means (for example, by isolating the most inflexible elements of such a stakeholder group and compromising with others).
10. Be your own harshest critic. You don't want to tear yourself down, but you are in a position to best understand the assumptions and attitudes on which your positions and suggestions are based, and hence to be able to challenge them most effectively. The point should not be mere self-skepticism or self-destruction, but continual challenge of what you believe. If you do this well, then you are far better placed to respond to external challenges and unjustified critiques.

Leadership, in the end, is not just a collection of characteristics. It involves personal growth, practice of many different skills, and continual learning. But perhaps most of all, it requires accepting the responsibility that a technologically trained professional has in today's society, dominated as it is by technology systems of rapidly growing complexity, with costs and benefits that are dramatic and uncertain. This book began with an observation from Goethe, and it appropriately ends with it, because it remains true now as it did many years ago when Goethe looked out at the then-difficult beginnings of modernity and wondered how we would cope with what we had already created:

He, only, merits freedom and existence
Who wins them each day anew.

—GOETHE, *FAUST*[2]

Exercises

15.1 Write imaginary résumés for yourself, five and ten years from now. What do they tell you about your goals?

15.2 Why is continual learning an important part of sustainable engineering and leadership?

[2]J. W. von. Goethe, *Faust Parts I and II*. trans. S. Atkins (Princeton, NJ: Princeton University Press, [1833] 1984).

APPENDIX A

Introduction to Sustainable Engineering: Generic Syllabus

INTRODUCTORY NOTES

1. The following generic syllabus is keyed to a semester system, but can be shrunk or expanded as necessary. Although it can be taught as a part of the core undergraduate (or graduate) engineering curriculum, there are often significant time and resource constraints associated with introducing major new material into that core. Accordingly, it may be preferable to teach this course as a "required distributional credit" course. Engineering students have few free credits, but are often required to take nonengineering courses in order to broaden their perspective; it is far preferable to have a course that provides them with relevant information about the context of modern engineering such as this, rather than one of little relevance, quickly forgotten. In this manner, the material can be introduced in a way that avoids conflict with the existing engineering curriculum and enhances the educational experience, rather than simply piling on more requirements.

2. The course is designed to be taught in two parts. One is a weekly lecture, which introduces each topic and can be done in large groups. Because the topics in these lectures vary widely, guest lecturers from other departments of the university or college might be necessary; alternatively, Arizona State University plans to offer a set of videos of the lectures. The second part is a smaller seminar, which allows students to participate in a more comfortable environment, encourages learning about effective communication because students cross-edit each other's essays, and provides a small environment for role-playing games and other activities that are not possible in a large lecture format. Seminar materials, including slide shows on relevant topics, are also available from Arizona State University. In this way, colleges can choose to either offer this sort of course on their own or with lectures provided via video and the seminars and grading done by their faculty.

3. The syllabus indicates the pedagogical goals of the course and how it fits with ABET requirements, as well as incidentals such as suggested grading criteria (ABET is the accrediting organization for engineering schools in the United States; other organizations may be involved in other countries). Obviously, this book is a stand-alone resource, and can be used in whatever way works best with an institution's culture and existing programs.

Nonetheless, the hope is that, by providing this material and guidelines, and associated support material, the introduction of sustainable engineering at a large scale, and the consequent innovation in engineering education, can be done as efficiently, rapidly, and effectively as possible. In particular, the syllabus, available from the author in electronic format, is designed so it can be used if desired virtually as is, with appropriate tweaks to reflect specific faculty information, and institutional practices and requirements.

4. This generic syllabus assumes the main audience for this book and course are undergraduate engineering students. This need not be the case and, indeed, mixed student populations are to be encouraged for two reasons. First, it is useful for engineering students to interact with those in other disciplines on issues involving technology and built systems, because it helps them understand how different people have differing perspectives on artifacts and technologies. Second, it is very useful for non-engineering students to understand more about technological innovation and the importance and power of technology systems. Depending on the goals of the institution, a good way to encourage such interaction and learning is to create mixed small seminars, which is where most of the interactive communication takes place.

AN INTRODUCTION TO SUSTAINABLE ENGINEERING: TECHNOLOGICAL, SOCIAL, AND SUSTAINABLE SYSTEMS

Lecture Hours and Location:

Small Seminar Hours and Locations:

Professor Contact Information:

Instructor Contact Information:

Course Objectives

To introduce students to the theory and practice of sustainable engineering, and the importance and role of technological, social, and sustainable systems in the modern world, which is increasingly characterized by integrated human/natural/built complex adaptive systems at local, regional, and global scales. Emphasis will be on characteristics and fundamentals of technology systems; complex adaptive systems behavior and evolution; and associated cultural, ethical, and managerial behaviors, with a focus on the need for multidisciplinary approaches and sustainable engineering; and current patterns in technological evolution. The course provides a framework for the theory and practice of sustainable engineering.

At the end of the course, students should be able to:

1. Explain the importance of technology and technological systems
2. Explain the social and environmental implications of design, construction, operation, and management of technology systems
3. Identify and explain critical principles of complexity and complex systems
4. Be able to use these concepts and principles to explore a topic of their choice in a systemic and integrated way in their term paper

In addition to domain-specific goals, this course is also intended to help students:

1. Learn to communicate in short essay form and in small group discussions
2. Understand issues and impacts associated with technology systems and emerging technologies at a broad cultural and geographic scale extending across urban, regional, national, and global scales
3. Understand the need and develop the capability to participate in lifelong learning

In terms of ABET criteria, the course will enable students to:

1. Understand professional and ethical issues in the context of engineered and earth systems
2. Learn to communicate in short essay form
3. Understand issues and impacts of engineering solutions at a broad cultural and geographic scale extending across urban, regional, national, and global scales
4. Understand the need and develop the capability to participate in lifelong learning
5. Take into consideration contemporary issues and environmental impacts in civil and environmental engineering practices

Introduction

The Industrial Revolution and continuing dramatic and accelerating changes in economic, technological, and cultural systems have fundamentally changed the way people live, relate to each other, and affect natural and built systems. In fact, many scientists are increasingly referring to our modern period era as the "Anthropocene," which can be roughly translated as the Age of the Human. Moreover, the accelerating pace of technological evolution—particularly the coming convergence of nanotechnology, biotechnology, robotics, information and communication technology (ICT), and cognitive sciences—will both reinforce the human domination of the dynamics of natural systems, and pose significant challenges to existing cultural and ethical norms and patterns. Thus, it is not possible to understand the modern world and to

make intelligent choices about the future without some understanding of technological systems and emerging technologies, and the complex systems of which they are a part.

Accordingly, this course will provide students with an introduction to technology and complex systems, and to sustainable engineering, and explore the relationship between technology and sustainability. It will also introduce students to the implications of understanding the Earth as a terraformed planet, and the technological, economic, and cultural patterns that have contributed to its evolution. The potential operational, cultural, and ethical implications of future evolutionary pathways will be explored, with emphasis on the challenges they pose and the role of technological systems in both creating and helping to address such challenges.

The course will consist of 1.5 hours of lecture and 1.5 hours of discussion in small groups with the teaching assistants (TAs), each week, except where otherwise noted. The course grade will consist of three components: class participation, a homework component that will primarily reflect short essays on class topics, and a term paper.

Homework will be a one-page, 1.5-line-spaced essay each week on the topic assigned for that week. It will be due the second class of each week at the beginning of class, where it will be collected by the TA and cross-edited by the students. The reading assignment for each week includes reviewing the slide package for the next week's class prior to class.

Grading

Class participation: 30 percent

- Lecture attendance and seminar attendance is mandatory; each unexcused failure to attend either is minus 2 percent of the class grade, up to the total of 30 percent
- Article presentation in seminar: 5 percent of the 30 percent

Homework: 35 percent

- Each missing weekly essay is minus 4 percent *from the overall class grade*
- Failure to meet minimum quality standards or a late paper is minus 2 percent *from the overall class grade* (up to a total of 35 percent)

Term paper: 35 percent, of which

- Acceptable topic submission: 5 percent
- Acceptable paper outline (schedule below): 5 percent
- Grasp and use of course concepts (e.g., complexity, technology systems): 25 percent
- Organization, structure, and persuasiveness: 25 percent

Proper format: 20 percent

Presentation (e.g., no typos, complete sentences): 20 percent

Term paper topic due:

Term paper outline due:

Term paper due:

Note: Term papers will be checked for plagiarism (copying language or barely modifying language that is not your own, without attribution). Plagiarism is a serious academic offense, and will result in a minimum of a grade of 0 (zero) for the term paper. If you have any questions about what constitutes plagiarism, contact your instructor or bring it up in your discussion group.

Primary Texts

1. Braden Allenby, *Theory and Practice of Sustainable Engineering*
2. Orson Scott Card, *Ender's Game,* or a similar science fiction novel (e.g., by William Gibson, Neal Stephenson, Vernor Vinge, or Charles Stross, depending on preference)
3. Samuel C. Florman, *The Existential Pleasure of Engineering*

Weekly Schedule (note that the topic of the week's lecture is given for each week; the second session of each week will consist of a small seminar where [a] the weekly essays will be collected by the TA and handed out randomly for cross-grading, with students commenting on style and substantive content of each essay in a mutually supportive learning environment; [b] additional open discussion of issues raised by the weekly lecture; and [c] additional instruction in writing, speaking, and learning skills). Slide packages for each lecture will be placed on Blackboard; each lecture will reference, but not necessarily go through, that week's set of slides.

WEEK 1: Welcome to the Anthropocene. A discussion of the increasing impact of humans on the world, including information on demographic trends (urbanization, population growth, per capita material, energy, and water consumption), global economic history, natural cycles (especially C and N), and how they interact in various systems, such as biofuels as response to global climate change.

Reading: Chapter 1, Allenby; slide package (in this and all subsequent weeks).

Homework (to be handed in at TA session on week 1): One-page essay from Chapter 1 exercises.

WEEK 2: Important Themes of the Human Earth. A number of general themes are important for understanding the world as it now exists and as providing

context for engineering and managing technology today. These include the criticality of technology as a contributor to and shaper of accelerating economic, environmental, social, and cultural evolution; the increasing information density and complexity of the world as many physical domains are redefined into information structures (e.g., genetics and bioengineering); and the growth of active information functionality within built environments (e.g., the "cognitive city" is a complex of "smart" materials, "smart" buildings, "smart" infrastructure, and "smart" integrated infrastructures). Moreover, natural systems are increasingly integrated with human and built systems, and thus become subject to their dynamics (e.g., reflexivity, intentionality); examples might include genetic engineering and its commoditization through intellectual property law, or carbon cycle and sulfur cycle management. Institutionally, the rapidly evolving technological frontier creates significant social pressures; fundamentalism as a response to modernity becomes increasingly powerful, and professionals and firms are being charged by society with responsibility not just for their actions, but for their technology systems (cf. Monsanto and genetically modified organisms). It thus becomes increasingly possible that technological evolution will become discontinuous in terms of cultural ability to adapt, especially because technological evolution continues to accelerate. From a cultural and political perspective, the end of the bipolar order imposed by the Cold War has destabilized global power relationships, resulting in cultural conflicts and, in some cases, terrorism. At the same time, the role of nation-state is changing profoundly, leading to more diffuse power structures, different and more varied forms of community (e.g., online social networks), and a loss of the traditional monopoly of the state on military power (e.g., technologies of mass destruction have been democratized).

Reading: Chapter 2, Allenby.

Homework: One-page essay chosen from Chapter 2 exercises.

WEEK 3: Complexity, Contingency, and Accelerating Change. There are very important differences between complex and simple systems. Simple systems are generally intuitively understandable, and their dynamics, such as "cause and effect," are relatively easy to understand. The anthropogenic Earth, however, is characterized by complexity, which can be thought of as including four forms of complexity: static complexity (number of components and links among them); dynamic complexity (introduced by features such as lag times and feedback loops that operate as a system moves through time); "wicked" complexity, which comes into play as humans and their institutions get involved; and scale complexity, as humans increasingly operate at the scale of regional and global natural and built systems. These operate together to create

radical contingency in the modern world, as it becomes difficult to determine what assumptions and institutions will remain valid over time.

Reading: Chapter 3, Allenby.

Homework: One-page essay chosen from Chapter 3.

WEEK 4: Sustainability. "Sustainability" was popularized as "sustainable development" by the Brundtland Commission in 1987, and has two basic themes: an emphasis on environmental quality, and a demand for increased equality in wealth distribution within and among generations. Despite its high popularity, it remains an ambiguous and somewhat contentious concept, and difficult to translate into operational terms that are easily implementable in the design and management of technology or earth systems. The sustainability dialog also tends to be naïve, if not somewhat skeptical, of engineering and technology, which leads to significant potential blind spots given the importance of emerging technologies to the shape and dynamics of future social, economic, and environmental systems. It is clear, however, that engineering and technology management today requires a greater focus on the environmental, social, and cultural dimensions of a design or project. Accordingly, much of sustainable engineering is, in fact, learning to translate the mythic language of sustainability into methods and inputs that can inform better engineering and technocratic decisions.

Reading: Chapter 4, Allenby.

Homework: Exercise 4.4, using memorandum format.

WEEK 5: Homo Faber: Human History and Technology. It is not surprising that periods of human development, such as "Neolithic" or "the Bronze Age," have referred to the dominant technologies of the time, because humans, their institutions, and their societies have always been coupled, and indeed to some extent defined, by their technologies. Indeed, since the Industrial Revolution economists have used the idea of "long waves" of innovation, characterized by particular technology clusters such as coal and the steam engine, or automobiles, to help understand not just technological evolution, but also economic, social, and cultural evolution as well.

Reading: Chapter 5, Allenby.

Homework: Essay from Chapter 5 exercises.

WEEK 6: Understanding Technology. Technologies have generic characteristics that tend to be fairly common. For example, they may grow vigorously when young, but they slow as they grow old; eventually, even the most fundamental infrastructures are replaced. They must be understood and framed as

systems if fundamental change is desired. For instance, if plug-in hybrids, and solar and wind renewables, are to be introduced into a developed economy, the timing will not be determined by how long it takes to build and design them, but rather how long it will take to rebuild the grid to handle greater variability of supply and a large demand spike. More broadly, many cultures have developed technologies, sometimes independently of each other, but few cultures have successfully made technology a core of their success. Part of the reason is that technology is not just widgets or software, but is also a social and cultural activity, and thus can be inhibited or encouraged by different cultural patterns.

Reading: Allenby, Chapter 6.

Homework: Essay from Chapter 6 exercises.

WEEK 7: Industrial Ecology. Industrial ecology, a relatively new field, is the objective, multidisciplinary study of industrial and economic systems and their linkages with fundamental natural systems. It incorporates, among other things, research involving energy supply and use, new materials, new technologies and technological systems, basic sciences, economics, law, management, and social sciences. Although still in the development stage, it provides the theoretical scientific basis upon which understanding and reasoned improvement of current practices can be based. Industrial ecology focuses on long-term habitability rather than short-term or ad hoc approaches, attempting to understand anthropogenic disruption to fundamental natural systems and cycles rather than just responding to localized perturbations. Typical industrial ecology approaches include mass-flow analysis to understand energy and material flows through economic and environmental systems, and the linkages among them; popular tools based on this approach include life-cycle analysis, or LCA. It is important to recognize that industrial ecology and associated practices in general continue to exhibit a strong bias towards the environmental domain, in part reflecting their origin in environmental engineering and science, and in part reflecting the complexity and strong normative dimensions of social issues, which make them much harder for engineers to quantify and evaluate. Industrial ecology thus forms an important base for sustainable engineering, but because it retains a focus on environmental, as opposed to social and cultural, domains, it cannot in itself be considered adequate for sustainable engineering.

Reading: Allenby, Chapter 7, first half of *Ender's Game*.

Homework: Essay from Chapter 7 exercises.

WEEK 8: The Five Horsemen: NBRIC. The railroad, a core technological system for a technology cluster, had major impacts even though it was just one technology system. We are currently seeing not one, but five, foundational

technology systems in a period of accelerating evolution: nanotechnology, biotechnology, robotics, information and communication technology, and cognitive science (NBRIC). These technologies in some ways are the logical end of the chapter of human history that began with the Greeks 2,500 years ago. Nanotechnology extends human will and design to the atomic level. Biotechnology extends it across the biosphere. ICT gives us the ability to create virtual worlds at will and facilitates a migration of functionality to information rather than physical structures. Robotics and biotechnology merge the biological and technological realms, enabling integration at the level of information systems. Cognitive sciences rationalize cognition and thus enable ever expanding cognitive networks that increasingly merge human and technology systems. Consider, for example, the way that Google so dramatically extends human memory, creating a cognitive system that includes not just the human elements, but vast swaths of Net technology; or the way that many militaries are building "augmented cognition" technologies, where technologies intended to scan the battlefield for threats are integrated into each soldier's cognitive systems. Current accelerating rates of technological evolution are not only unprecedented; they have the effect of dramatically extending the spaces within which humans can, intentionally and unintentionally, impact existing systems and design new ones. In doing so, these rates not only raise the level of complexity of systems that we must strive to understand. Because they also give us rapidly increasing tools to design the human itself, they render contingent much of what we have taken to be fixed.

Reading: Allenby, Chapter 8; second half of *Ender's Game*.

Homework: Essay 8.4 from Chapter 8 exercises.

WEEK 9: Case Studies in Technology Systems: Green Chemistry. This is the first case study of a major technology system—chemistry—which is being modified to reflect at least some of the concerns behind sustainability. It is apparent that each level within a particular technology offers its own unique challenges and opportunities for the sustainable engineer or technology manager. Manufacturing a pharmaceutical that is metabolized in the body has different social and environmental issues associated with it than designing an antimicrobial soap with active ingredients that survive sewage treatment technologies and become part of surface water chemistry. It is also apparent, however, that each technology, as a system, is more complex and integrated than engineers, businesspeople, and policy makers realize. This marks an important area where sustainable engineering offers insights that can improve professional performance significantly.

Reading: Allenby, Chapter 9.

Homework: Essay 9.5, Chapter 9.

WEEK 10: Case Studies in Technology Systems: Information and Communication Technology (ICT). Many of the most environmentally and socially important technological changes characterizing the Anthropocene arise directly or indirectly from the rapid evolution of information and communication technologies, in part because the services involved, from television to social networking sites, directly impact social behaviors and norms in predictable and unpredictable ways. Virtual reality and gaming technologies, for example, are becoming increasingly attractive as physics engines and display technology become more powerful; social networking sites are increasingly substituting for what traditionally were face-to-face interactions. Because these changes are so new, they are not well understood even as the underlying technology platforms accelerate in power and performance.

Reading: Allenby, Chapter 10.

Homework: Essay from Chapter 10 exercises.

WEEK 11: The Five Horsemen, Military Operations, and National Security. Throughout history, technological evolution and military activity have been linked. The existential challenge to society represented by warfare, combined with the immediate advantage that new technology can deliver, tends to accelerate technological innovation and diffusion. The relationships between the resulting technology systems, and consequent social and ethical issues and changes, are quite complex, however, and understanding and managing them to enhance long-term military advantage and security is a critical and underappreciated challenge. This is particularly true when, as now, technological change is both rapid and accelerating, posing the risk of cultural backlashes that could affect both short-term mission capabilities and longer-term security interests. Moreover, many technologies of sufficient power to be of interest militarily have at least the potential to be deeply destabilizing to existing economic, social, and technological systems. Examples might include the possibility that military radio frequency identification (RFID) sensor systems, insect robots, and cyborgs are shifted from theater intelligence to domestic intelligence; that telepathic helmet technology transitions from a small unit communication enhancement to a nonintrusive thought detection device in civil society; or that warrior enhancement technology results in radical life extension for selected civilian populations. Emerging technologies are likely to have similar destabilizing effects within the military, potentially affecting not just military operations, but military culture and organization, as well as broader social perspectives on military initiatives generally.

Reading: Allenby, Chapter 11.

Homework: Essay from Chapter 11 exercises.

WEEK 12: The Engineer as Ethicist: Macroethics. From the viewpoint of the engineer and the technologist, four important characteristics of the Anthropocene differentiate it from the traditional human systems within which existing ethical structures have developed. The first is that the earth systems characteristics of the Anthropocene are neither "human" nor "natural," but highly integrated composites of both. The second is that, as a result, the dynamics of such anthropogenic systems include the reflexivity and thus unpredictability of human systems. Third, these systems are highly interconnected. Managing global climate change is difficult precisely because the climate system is tightly coupled to human economic and technological systems and their future paths, to powerful cultural and ideological systems, and to other natural systems such as the carbon and nitrogen cycles. Finally, technology systems are major vehicles by which this complexity, integration, and unpredictability are created. Existing ethical systems and many proposed principles such as the Precautionary Principle (don't implement a technology until you are sure the risks it poses will be less than the existing risks) are inadequate to this level of unpredictability, requiring the sustainable engineer to be much more sophisticated regarding the culture, ethical systems, and contingency of the frameworks within which he or she operates.

Reading: Allenby, Chapter 12; watch *The Matrix* (first film in series); Florman, Chapters 1–3.

Homework: Essay: How do you know either reality in *The Matrix* is more "real" than the other? (Is there a purple pill? A puce one? Why not?) And if even someone in those realities can't tell which is which, what do you think of the morality of Neo's decision to destroy the Matrix?

WEEK 13: Sustainable Engineering in the Anthropocene: The Aral Sea, the Everglades, and Adaptive Management. The challenge of managing large resource systems, such as fisheries and forests, to be sustainable over time is substantial. In recent years, the difficulty of successfully managing regional resources, such as the Great Lakes, the Aral Sea, or the Everglades, has also become apparent as pressures on them have grown with increasing population and economic exploitation. Designing and managing the Everglades, for example, raises complex problems arising from mutually exclusive stakeholder value systems in the context of a highly valued, unpredictable, and complex resource regime. "Adaptive management," a process with some similarity to industrial ecology and systems engineering, has been developed primarily by ecologists for design and management of complex resource regimes such as fisheries, forests, and watersheds, and should be part of the toolbox of the sustainable engineer.

Reading: Allenby, Chapter 13; Florman, Chapters 4–6.

Homework: Essay from Chapter 13 exercises.

WEEK 14: Earth Systems Engineering and Management (ESEM): Sustainable Engineering at a Planetary Scale. Although still a very young area of study, sustainable engineering at the ESEM level can draw on a number of areas of study and practice that are already fairly robust, including industrial ecology, adaptive management, and systems engineering. In addition, there is a growing literature on industrial management and complexity that is quite useful, in part because both business managers and engineers face a similar type of challenge: creating solutions that work in the real world for complex problems. Further, sustainable engineering can draw on experience from a number of other fields of study, including urban and regional planning, sociology and the social construction literature, and the management of complex technological systems such as the Internet, global air transport and financial networks, military weapon systems, and the like to develop a set of initial ESEM principles.

Reading: Allenby, Chapter 14.

Homework: Chapter 14, Exercise 14.3.

WEEK 15: The Engineer as Leader. Sustainable engineering requires many things of professionals: commitment; respect for values and opinions that differ among themselves and from the ones we may hold; a willingness to understand and work with social, cultural, and environmental contexts. But it also requires that, as knowledgeable citizens in an increasingly technological world, engineers function as leaders within their institutions, communities, and society at large.

Reading: Allenby, Chapter 15; Florman, Chapters 7–11.

Homework: Essay from Chapter 15 exercises.

APPENDIX B

Glossary of Key Terms

Activity fulfillment: A transportation engineering concept, *activity fulfillment* is a guide to design based on enabling people to perform the activities they want in as efficient a way as possible. Although in traditional transport engineering activity fulfillment is a way to compare different transportation infrastructure designs, this concept can be generalized to include "virtual activity fulfillment" through, for example, substitution of information and communication technology options for physical transportation.

Adaptive management: Growing primarily out of the ecological literature, *adaptive management* is a multidisciplinary approach to complex ecosystems, such as the Florida Everglades or Baltic Sea, which emphasizes continuing learning and adaptation of policy and management to changes in the system of interest, based on new data and emergent behaviors.

American exceptionalism: An American doctrine holding that the United States is qualitatively different from, and better than, other countries because of its history and emphasis on individual freedom and liberty. In this view, America is seen as the "shining city on a hill" serving as a model for other countries to aspire to. The "shining hill" language has biblical roots, but in its American exceptionalism application originated with the Puritans. It has been used by American presidents such as John F. Kennedy and Ronald Reagan, and the doctrine remains very much alive today among some American political groups.

Anthropocene: A term increasingly used by scientists and others to refer to the modern era; it translates roughly as the "Age of the Human," and is intended to indicate that the role of humans is the dominant characteristic of this period of the planet's history.

Anthropogenic Earth: Literally, a "human-made" Earth; the term refers to the ubiquitous impacts of human activity on the planet and the rise of regional and global integrated human/built/natural complex adaptive systems characteristic of the current era. The term does not imply that humans deliberately have designed earth systems, many of which are too complex to be understood to that degree, but rather reflects the comprehensive impacts of human activities and technologies.

Augmented cognition: *Augmented cognition*, or *augcog*, is a research and development field that involves the diffusion of cognitive function across integrated human–technology networks. For example, the ability of a search engine to significantly enhance internal human memory function is a form of augcog.

Carbon determinism: *Carbon determinism* is the belief that global climate change is the determining factor in the evolutionary paths of human cultures, societies, and technologies. It is somewhat different from traditional determinism theories such as economic determinism (q.v.) in that it is usually intended as prescriptive rather than descriptive (global climate change *should* determine future political and social trends, rather than holding that it already *has*). As a concept, it has the same flaws as all determinisms do: no single factor or perspective is adequate to explain the behavior of complex adaptive systems, such as human societies.

Carbon footprint: A *carbon footprint* is the amount of carbon dioxide (sometimes expanded to include all greenhouse gases) required to produce a particular product, service, or device. There is no generally accepted methodology for deriving carbon footprints, and different methods give very different results, so any specific result should be treated with great caution.

Carbon fundamentalism: The view that many if not most human activities (owning pets, having children, travel) can and should be evaluated primarily by their impact on global climate change, usually measured by releases of carbon dioxide and other greenhouse gases.

Climate change: *Climate change* refers to long-term shifts in regional and global weather patterns. It cannot be perceived through human senses, and is therefore identified using large data sets and computer models. It is not the same as "global warming," which is generally used to refer to climate change caused by human activity, especially release of greenhouse gases to the atmosphere.

Complex adaptive system: *Complex adaptive systems*, or CASs, are dynamic, interacting, self-organizing networks that change in response to shifts in internal or external states, frequently displaying emergent behaviors, which are unpredictable and arise not from individual components of the CAS, but from the CAS as a whole. Examples of CASs might include stock markets, large economies, political systems, ecosystems, and the human brain.

Computer–Brain Interface: *Computer–brain interfaces*, or CBIs, are devices that allow connection between human brains and technology systems. Commercial devices already enable some control of avatars, or balls floating on air columns, through primitive CBIs (specially designed headbands); advanced CBI devices include chips embedded in monkey brains, which allow the monkey to control running robots or detached mechanical arms many hundreds of miles away.

Dead Zone: *Dead zones* refer to areas that appear on the oceanic floor in relatively shallow estuarine and continental shelf areas when phosphorous and nitrogen from land-based activities, primarily agricultural, flow into oceanic areas and cause algal blooms. The algae and small organisms that feed on the

algal bloom then die and sink to the bottom, where their decomposition uses all available oxygen, creating a dead zone. The best known dead zone is in the Gulf of Mexico.

Disruptive technology: A *disruptive technology* is one that unexpectedly and fundamentally changes existing markets or the viability of current technologies. For vacuum tube manufacturers, solid state electronics was a disruptive technology; for telephone companies, the shift from dedicated point-to-point networks, to telephony based on Internet platforms, was a disruptive technology.

Dynamic complexity: *Dynamic complexity* arises as systems iterate, generally by moving through time, and change internally in response to changes in external and internal states. Dynamic complexity does not require static complexity (q.v.); even systems that are very simple when static can exhibit dynamic complexity under certain circumstances.

Economic determinism: *Economic determinism* is the belief that economic factors explain, and in fact determine, the paths of human history and culture. Economic determinism is often attributed to Karl Marx, who emphasized the importance of economic institutions in shaping political and historical trends. Like any determinism, the main problem with economic determinism is that no single factor or perspective is adequate to explain the behavior of complex adaptive systems, such as human societies.

Ecotoxicity: The inherent ability of a material to cause harm to a target ecosystem.

EIO-LCA: *Economic input–output life cycle assessment*, or EIO-LCA, is a life cycle assessment technology developed by the Carnegie Mellon University Green Design Institute. It combines the economic input–output matrix method with an environmental component to produce top-down analyses of environmental impact of changes in product, process, or material flows in various sectors (available at www.eiolca.net).

Embedded water: The amount of water used to produce, package, and transport a product or material, as distinguished from the water actually contained in the product. A gallon of milk contains a little less than a gallon of water, but the embedded water content is far larger, because cows drink water and eat grass, which requires water to grow, to produce the milk.

Emergent behavior (or characteristics): as complex systems self-organize and evolve, they produce unpredictable and novel features and behaviors at the level of the system itself, which could not be predicted a priori from knowing the nodes and connections that compose the system.

Engineering: *Engineering* is structured, usually quantitative, problem solving, a set of disciplines and practices that apply scientific and mathematical principles and knowledge to design, construct, operate, and maintain the built

environment, including everything from the familiar machines and infrastructure that surrounds us, to the software and information systems that increasingly provide services to humanity.

Enlightenment: The *Enlightenment* is the period in European history around the seventeenth and eighteenth centuries when reason, especially reason supported by scientific experimentation and observation, replaced religion as the source of authority and the primary mechanism for exploring human and natural systems. It is also known as the Age of Reason.

Environmental footprint: The overall environmental impact associated with a product, service, activity, or institution. Methods for calculating environmental footprints vary greatly, and some critics argue that all of them have been designed to overstate environmental impacts.

Five Horsemen: A term referring to the five foundational emerging technologies: nanotechnology, biotechnology, robotics, information and communication technology, and applied cognitive science. Because these technologies are not just rapidly changing, but are foundational for many other technologies, and are combining in unpredictable and powerful ways, they actually represent rapid and unpredictable evolution across the technological frontier as a whole.

GDP: *GDP* stands for "gross domestic product," which is the market value of the goods and services produced in a particular country in a year.

Geoengineering: *Geoengineering* is a term of art referring to development and deployment of technologies at large scale to reduce global climate change forcing. Two classes of technology are recognized in the geoengineering literature: carbon dioxide removal (CDR) techniques, which though biological or mechanical means remove CO_2 from the atmosphere; and solar radiation management (SRM) techniques, which reflect incoming solar radiation back into space before it has an opportunity to directly or indirectly warm the atmosphere.

Global warming: *Global warming* is generally used to refer to a warming of the atmosphere induced by increased human emissions of gases that capture radiation from the Earth, especially carbon dioxide and methane. The degree to which humans are responsible for global warming (sometimes known as "anthropogenic climate change") is politically contested. "Global climate change" is a preferable term, because even if the atmosphere as a whole warms, climate in particular areas may become either warmer or cooler, depending on local and regional dynamics.

GMOs: *GMO* stands for "genetically modified organism," and is almost always applied in practice to agricultural crops that have been genetically engineered to exhibit desirable traits, such as corn or soya that resist insect infestation.

Golden Age: A hypothetical period in the past when humans were in harmony with each other and their environment. A common manifestation of this

line of thinking is the idea, widespread among some sustainability advocates, that more primitive communities are necessarily more sustainable, a wistful belief that research has found to be predominantly ideological rather than factual.

Great Chain of Being: The *Great Chain of Being* is a conceptual framework of the Universe perfected in the Christian medieval period in Europe that differentiates plants, beasts, and humans based on the functionality exhibited by their souls—plants having less than beasts, which in turn have less than humans. It defines a structured hierarchy, with pure spirit and perfection (God) at the top, and pure matter and imperfection (rocks and other materials) at the bottom. In between, in order, come all things; angels are next to God, while plants, then beasts, are above matter. Humans are at the fulcrum point, for they are both spirit and matter. The Great Chain of Being served to justify the strong hierarchical structures that characterized society and the Church at that time. Because it is ordained by God and not a matter of choice, changing it by, e.g., technological means is a violation of the basic order of the Universe and God's will, and is thus blasphemous.

Great Divergence: For most of human history, almost everyone lived at a subsistence level. With the Industrial Revolution, however, a large and growing gap opened between the average income of individuals in countries with advanced economies and countries that were less economically developed. This gap is known as the "Great Divergence."

Greenhouse effect: The *greenhouse effect* is a process where infrared radiation from a planet's surface is absorbed by gases in its atmosphere—such as water vapor, carbon dioxide, and methane in the case of the Earth—thereby raising the temperature of the atmosphere as a whole. Many people incorrectly equate the greenhouse effect, a scientifically proven process, with "climate change" (q.v.).

Hierarchy of control: An occupational safety and health principle that ranks the effectiveness of interventions to protect workers. The most effective mechanism is to eliminate the source of hazard, followed by engineering controls that reduce the hazard; the least effective technique, because it is relatively easy to evade, is to introduce personal protective equipment.

High modernism: In literature and the arts, *high modernism* refers to the most developed modern aesthetic, as in "high Renaissance." In technology and science studies, however, the term is generally pejorative, and implies a hubristic and elitist elevation of engineering, design, and scientific principles over human interests, usually supported by an authoritarian polity.

ICT: *ICT* stands for "information and communication technology," a broad category covering both content generation and information technology, including computer, smart phone, consumer electronics, and transmission

technologies; the Internet in its various guises; storage and entertainment media; software packages and services such as social networking services, network security packages, and search engines; and much more.

Industrial ecology: The study of industrial and economic systems and their linkages with fundamental natural systems, usually with an emphasis on environmental considerations, using tools such as life cycle assessment (LCA) and mass flow analysis (MFA).

Legacy system: A *legacy system* is one that is already in place and must be considered as further designs are implemented. Thus, for example, building the smart grid in the United States cannot be done without taking into account the existing power plants, distribution systems, legal and regulatory arrangements, and similar legacy systems that already exist. Not only do legacy systems complicate design but, because they greatly complicate upgrades and changes, they can provide competitive advantage to newcomers who do not need to make accommodations for them.

Linear and nonlinear behavior: A system is said to be acting in its "linear regime" when it changes predictably and smoothly in response to changes in its environment, but in a "nonlinear regime" when it changes unpredictably and dramatically in response to a small change. The same system may behave linearly at some times, and non-linearly in others: a salt marsh may display gradual changes as it receives increasing loads of pollution (linear), but at some point may shift dramatically even with a small additional increment as its internal structure is unable to continue to absorb the pollution (a non-linear response; such dynamics are reflected in sayings such as "the straw that broke the camel's back").

Lock in: *Lock in* refers to situations where technological, economic, social, political, or other factors make changing to alternative systems much more expensive or difficult than would otherwise be the case. Technological lock in occurs where a particular technology is embedded in other technology systems so that its alteration is impeded: an example would be the way current energy infrastructure design and commercial production capability locks in gasoline engines in cars. Vendor lock in occurs where commercial arrangements or proprietary systems make using alternative vendors much more difficult. Social and political lock in occurs where change is impeded because strong social and political forces have grown up around a particular option. Thus, although the Kyoto Protocol approach to climate change has been relatively unsuccessful, the amount of energy, and institutional and psychological commitment already devoted by many people and organizations to that process means that it is strongly resistant to fundamental change.

Manifest Destiny: An American political doctrine popularized in the 1840s holding that the United States was intended by God to extend from the Atlantic to the Pacific oceans (q.v. American exceptionalism).

Master equation: In industrial ecology, the *master equation* qualitatively breaks down human environmental impact into three components: population level, consumption per unit population, and impact per unit consumption. For purposes of sustainable engineering, the equation has been redefined in this text to encompass sustainability, rather than just environmental, values.

Military determinism: *Military determinism* is the belief that war and military activities are determinative of social, cultural, and historical evolution. Despite the obvious impacts that war often has on the course of history, any such determinism is inadequate because no single factor or perspective is adequate to explain the behavior of complex adaptive systems such as human societies.

Nation-state: A "state" is a geopolitical entity (a "country") such as China, the United States, or Vietnam, while a "nation" is a cultural or ethnic entity, such as a Native American nation. A *nation-state* is therefore a geopolitical entity with a unified cultural or ethnic identity. The old Soviet Union, for example, was a combination of different cultures, and was therefore a state but not a nation, while the results of the breakup, such as Georgia, are nation-states. In practice, political boundaries seldom if ever neatly match cultural and ethnic boundaries.

NGO: A *NGO* is a nongovernmental organization, such as a human rights or environmental organization.

Orgware: *Orgware* refers to the institutional and organizational components of systems; compare with "hardware" for physical artifacts and "software" for computer code and information components.

Precautionary Principle: The *Precautionary Principle* holds that no new technology should be introduced unless it can be demonstrated to be less risky than the status quo (in the words of the UN World Charter for Nature, "where potential adverse effects are not fully understood, the activities should not proceed."). Taken literally, the Principle would not allow any more than trivial technology to be introduced, as the full implications of a technology cannot be understood except in real time as it is absorbed, and changed, by social, economic, and cultural systems.

Quad: A *quad* is 10^{15} BTUs; a BTU is a British thermal unit, which is the amount of heat required to raise the temperature of a pound of water one degree Fahrenheit.

Real time technology and policy assessment: Real time technology and policy assessment, or RTTPA, identifies a set of practices that enable more effective and rational engagement with complex adaptive technology systems by, among other things, continuing to interact with the relevant systems as they evolve in real time and developing options for responding to unpredictable developments through such techniques as structured scenario exercises.

Reductionism: *Reductionism* is an approach to understanding systems by breaking them down into simple parts and relationships that can be more easily analyzed. It is key to scientific analysis. For example, an environmental scientist might seek to understand the impact of lead on salmon by placing salmon into tanks, a controlled environment, and varying only the lead level in the water. Reductionism continues to be highly useful in scientific and engineering studies, but has the major limitation that it cannot explain emergent behaviors (q.v.) of complex systems.

Reserves: *Reserves* is a general term for bodies of materials such as ore, oil, or coal that have been identified and verified, and are probably recoverable using technology and assuming current economic and regulatory conditions. In some cases, categories of reserves such as "proven reserves" are defined not by scientific or engineering standards, but by financial regulatory bodies such as the American Securities and Exchange Commission, in order to avoid securities fraud (e.g., firms claiming to own more recoverable minerals assets than they in fact do).

Resources: *Resources* are bodies of materials that are potentially recoverable, but that have not yet been fully characterized or evaluated. For virtually all materials, the resource number is far higher than the "reserve" or "proved reserve" number.

Resource base: The *resource base* is defined as all of a mineral commodity contained in the Earth's crust. For virtually all minerals, this number is unknown and certainly far larger than any other reserve or resource category.

S curve: Many natural and engineered systems display a characteristic life-cycle behavior where they grow slowly at first, then accelerate to an exponential phase, then slow down and eventually level out. The shape formed when this process is graphed against time looks somewhat like an extended "S," which explains the term. See Box 1.2 in the text for examples and further explanation.

Scenario planning: *Scenario planning*, or *scenario analysis,* is a technique of exploring unpredictable future events by creating and evaluating simplified alternative scenarios of what might occur. It is used by organizations such as militaries and corporations to practice adaptive responses rather than to try to predict actual futures (q.v. real time technology and policy assessment).

Smart grid: *Smart grid* is a somewhat informal term referring to an electric grid that uses digital, computer, and Internet technology to optimize resiliency, real-time balancing of supply and demand at local and regional scales (including, for example, unpredictably variable supply sources such as solar and wind power), and active management of appliances and users to optimize grid performance.

Software: Technically, *software* is the set of programs and data that run on a computer. The term is sometimes generalized to the information and

algorithms that run on any information processing system (e.g., "the human brain's software"). It is differentiated from "wetware"—biological carbon-based protoplasmic systems—and "hardware"—built artifacts such as chips, routers, and servers that run software.

Static complexity: *Static complexity* arises from the number of things and their connections (nodes and links) that are involved in a particular system. For example, an automobile or airplane simply sitting on the ground is a highly complex object in part because of all the parts, subassemblies, and physical and information linkages from which it is composed.

Sustainable chemistry: Chemical engineering, manufacturing, operations, and design that take appropriate account of environmental and social considerations over the life cycle of the chemical or material and related processes and products.

Sustainable development (sustainability): The traditional definition of *sustainable development* is from the World Commission on Environment and Development (also known as the Brundtland Commission) report, *Our Common Future*; it requires development that "meets the needs of the present without compromising the ability of future generations to meet their own needs."[1] Although the term has become fairly ambiguous due to overuse, it is generally held to require two core values: environmental quality, and equality within and among human generations.

Synthetic biology: *Synthetic biology* merges the theories, practices, and methods of engineering and biology to create biological forms and organisms such as agricultural GMOs (q.v.) that are not products of natural evolution by, for example, creating standard biological components that can be mixed and matched in target organisms to provide desired functions.

Systems: *Systems* are groups or networks of interacting, interdependent parts linked together by exchanges of energy, matter, and/or information. **Simple systems** are ones where changes in system output can be easily predicted by changes in input and system responses are linear: if you add sugar to lemonade, it will get sweeter in direct proportion to the amount of sugar you add. **Complex systems** are ones where output changes unpredictably as input changes, and are characterized by such phenomenon as threshold effects, lag times, lack of causal connections between input and output, and wicked complexity (q.v.).

Technological determinism: A belief that technology determines the course of history and culture. Analogous forms of determinism include the belief that economic forces determine the course of history and culture ("economic

[1] *Our Common Future* (Oxford: Oxford University Press, 1987), p. 8.

determinism," q.v., often associated with Marxism), or that environmental factors are similarly determinative ("environmental determinism"), or that anthropogenic climate change is so important and omnipresent that it is determinative ("carbon determinism," q.v.). The main flaw in any determinism is that no single factor or perspective is adequate to explain the behavior of complex adaptive systems, such as human societies.

Toxicity: *Toxicity* is the inherent ability of a material to cause harm to a target organism, organ, or biological system.

Triple Bottom Line: The *triple bottom line,* or TBL, is an extension of the more familiar corporate "bottom line" concept regarding economic performance; it suggests that the firm or organization should optimize not just economic performance, but environmental and social performance as well. Some have suggested a "quadruple bottom line," with cultural being added to the domains to be optimized, but that term is far less frequently used.

Wetware: *Wetware* is a term used to refer to biological systems as opposed to hardware or software systems. The human brain and nervous system, and the carbon-based living tissue that makes up plants, animals, and humans, are wetware.

Wicked complexity: *Wicked complexity* is a term popularized by H. Rittel and M. Webber; for sustainable engineering, it refers to complexity that arises when built and natural systems become integrated with human social, cultural, and economic systems, and thus display the reflexivity, unpredictability, and uncertainty characteristic of human systems. Importantly, systems displaying wicked complexity cannot be optimized, meaning that many of the tools that technocrats might use, such as operations research methods, are inapplicable.

X-ware: *X-ware* refers to system components that are not specifically captured by the usual terms referring to categories of system components, including hardware, software, wetware, and orgware.

Index

Page numbers followed by *f* and *b* indicate figure and boxes respectively.

A

Activity fulfillment, 77
Adaptive management, 342–345
  Aral Sea, 345–350
  defined, 344
  Florida Everglades, 350–355
  principles, 357–359
  themes, 355–357
Africa, age distribution in, 81, 83*f*
Age distribution, 78–85
  in Africa, 81, 83*f*
  in Egypt, 81, 82*f*
  in Europe, 78–80, 79*f*–80*f*, 81
  implications of, 81, 82
  life expectancy, 82, 83–84, 84*f*
  in United States, 81, 81*f*
Age of Reason, 56
Aging population, Applied ethics, 84–85
Agriculture, 265, 342–343
  in Aral Sea, 346
  for biomass, 43*b*
  employment shift for, 176, 177*f*
  evolution and, 21
  GMO in, 59, 63*b*–64*b*
  green revolution in, 20*b*
  in Southern Florida, 350, 354
  technologies for, 177
  water for, 32
AI. *See* Artificial intelligence (AI)
Airplane
  complexity, 101
  environmental impacts of, 92
  technology, 178–179
Ak-kum desert, 346
American exceptionalism, 305*b*
American Forest Congress, 151
Amu Dar'ya, 346
Analogies, and metaphors, 234*b*
Anthropocene, 18–19
  demographic trends. *See* Population
  globalization, 51–62
  ICT, 62, 65–67
  multiculturism, 51–62
  technology, 67–69
Antimicrobials, 256–261
Apollo spacecraft, 229
Applied ethics, 327. *See also* Ethics
Aral Sea, 345–350
  agriculture in, 346
  climate change impact on, 346–348
  inflow to, 346
  location of, 345
  population of, 346
  rivers, 346
  shrinkage of, 346, 347*f*
  as unintended impacts, 349
Aristotle, 70
Aristotle's De Anima, 218*b*
Arsenic, 251*b*–252*b*
  use of, 124
Artificial intelligence (AI), 238
AT&T, 202, 231, 396, 397
  telework program of, 285, 286, 287
Aucog. *See* Augmented cognition
Augmented cognition, 238
Automobiles, 40, 156, 159*b*, 237
  plug-in, 104*b*–105*b*
  Tata Nano, 91, 132, 134
  virtualization in, 170
Automotive technology, 237
Autonomic city, case study, 76*b*

B

Babbage, Charles, 93
Bell Laboratory, 113, 397
Bhagavad Gita, 155, 301
Bicycles, in urban environments, 75, 77
Biodiversity, 224*b*–225*b*
Biofuels, 104*b*
Biomass, 43*b*
Biomass Research and Development Board (BRDE), 42*b*
Biotechnology, 221–223
Birth control pill, 59
Blue baby syndrome, 39–40
Blue-collor workers, 153
Boston Central Artery/Tunnel (CA/T) project, 369
Bottom-up approach, 203
Brain, and machine, 228*b*–229*b*
BRDE. *See* Biomass Research and Development Board (BRDE)

C

Capitalism
  communism and, 106–107
  railroads and, 152–153
Caravel, 53
Carbon capture and sequestration (CCS), 363
Carbon cycle, 36–38. *See also* Climate change
  for bio-based products, 44*f*
  illustration, 37*f*
  management, 38, 38*f*
Carbon dioxide emission, 364. *See also* Climate change

Carbon footprint, 38
Carbon fundamentalism, 143*b*
Carnegie Mellon University Green Design Institute, 203
CBI. *See* Computer-brain interface (CBI)
CCA. *See* Chromated copper arsenate (CCA)
CCS. *See* Carbon capture and sequestration (CCS)
CERCLA. *See* Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA)
Chariot war, 301
Cheng Ho, 53
China, 18, 58, 343
  culture of, 330
  economy, 106
  irrigation in, 348
  leadership, 392
  mobile telephone subscription in, 172, 172*f*
  ship technology in, 53, 54
  water system in, 342
Chlorofluorocarbon (CFC), 249*b*
  case study, 253–256
  restriction on use of, 387–388
Chromated copper arsenate (CCA), 124
Cities. *See* Urbanization
Cities in Civilization, 368–369
Climate change, 335, 336
  and Aral Sea, 346–348
  carbon cycle, 36–38
  debates on, 35–36
  defined, 35
  in developed countries, 132, 134
  DFE and, 165
  geoengineering and. *See* Geoengineering
  model/modeling of, 108, 109, 110
  moral hazard, 364
  nitrogen cycle, 39–40, 39*f*
  physics of, 36*f*
Clouds, 364
Code of ethics, 331–332
Cognitive science, 238
Cogsci. *See* Cognitive science
Collapse, 143*b*
Commoditization, of landscapes, 235–236
Communism, and capitalism, 106–107
Complex adaptive systems, 43*b*
Complexity, 91
  challenges of, 91
  dynamic, 101, 102
  ICT and, 107
  model/modeling in, 107–111
  static, 101, 101*f*
  understanding, 106–107
  wicked, 102–103
Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), 245
Computer-brain interface (CBI), 227, 316
Confucian cultures, 330
Conservation biology, 224*b*–225*b*
Construction materials, 122
Corbusier, Le, 367
Core markets, 171
Corn-based ethanol, 42*b*–45*b,* 365
Counter Rocket Artillery Mortar system (CRAM), 320
Covering fire, 394
CRAM. *See* Counter Rocket Artillery Mortar system (CRAM)
Creosote, 152
Cronon, William, 73
Cultural conflict, 58
Cultures, and ethics, 330–331
Curators aquarum, 343
Cyborg insects, 310, 313–315
  policy response matrix, 314*f*

D

DDOS. *See* Distributed denial of service (DDOS)
DDT. *See* Dichlorodiphenyltrichloroethane (DDT)
Dead zone, 40
Descriptive ethics, 327. *See also* Ethics
Democratic Republic of the Congo (DRC), 171
Demographic trends, 69–87. *See also* Population
Design for assembly (DFA), 164–165
Design for disassembly (DFD), 165
Design for environment (DFE), 165, 272, 273*b*–275*b*
Design for manufacturability (DFM), 165–166
Design for regulatory compliance (DFRC), 166
Design for reliability (DFR), 166
Design for service and maintenance (DFSM), 166
Design for sustainability (DFS), 166
Design for testability (DFT), 167
DFA. *See* Design for assembly (DFA)
DFD. *See* Design for disassembly (DFD)
DFE. *See* Design for environment (DFE)
DFM. *See* Design for manufacturability (DFM)

DFR. *See* Design for reliability (DFR)
DFRC. *See* Design for regulatory compliance (DFRC)
DFS. *See* Design for sustainability (DFS)
DFSM. *See* Design for service and maintenance (DFSM)
DFT. *See* Design for testability (DFT)
Diamond, Jared, 143*b*
Dichlorodiphenyltrichloroethane (DDT), 260
Digital cameras, 174
Digital natives, 266. *See also* Information and communication technology (ICT); Virtual office
Digital nomads, 280
Digital photography, 174
Dike, Hoover, 351
Disruptive technologies, 176
Distributed denial of service (DDOS), 74, 309
DRC. *See* Democratic Republic of the Congo (DRC)
Dry cleaning, 20
Duty ethics, 332*f*, 333
Dynamic complexity, 101, 102

E

Earth systems engineering and management (ESEM), 361–388
  design and management principles, 381–387
  geoengineering, 363–366
  governance principles, 375–381
  high modernism, 367–370
  overview, 361–363
  theoretical principles, 370–375
Economic determinism, 143
Economic Input-Output Life Cycle Assessment (EIO-LCA), 203
The Economist, 266
Ecotoxicity, 133. *See also* Toxicity
Egypt, age distribution in, 81, 82*f*
EIO-LCA. *See* Economic Input-Output Life Cycle Assessment (EIO-LCA)
Elements, human mobilization of, 40, 41*t*
Embedded water, 32–33, 33*f*, 34*f*
Emergent behavior, 44*b*
Employment, technological impact on, 176–177, 177*f*
Energy, 26–31
  consumption patterns, 27–31, 28*f*–30*f*
  global consumption over the past 154 years, 27*f*
Energy materials, 122
Energy resources, 127, 127*f*
Engineering, 15–16
  defined, 16
  sustainable. *See* Sustainable engineering
Enlightenment, 55–56
Enron, 395
Environmental footprints, 42
Environmental Protection Agency (EPA), 124
EPA. *See* Environmental Protection Agency (EPA)
Ethical codes, 331–332
Ethics
  defined, 327
  framing, 328–335
  law and, 327–328
  standards, 328
Europe, age distribution in, 78–80, 79*f*–80*f*, 81
Everglades, Southern Florida, 102, 350–355
  major zones, 352*f*
  nesting in, 354
  restoration project, 354–355
  water flow/quality in, 353–354, 353*f*
Everglades National Park, 351

F

Facebook, 66, 232, 266, 296. *See also* Social networking
Firms. *See* Network-centric firms
Five Horsemen
  biotechnology, 221–223
  concept, 215–216
  ICT. *See* Information and communication technology (ICT)
  nanotechnology, 219–220, 221*f*
  robotics, 223, 226–227, 226*f*
Florida
  Everglades, 102, 350–355
  population growth in, 350
  satellite photograph of, 351*f*
Food and Drug Administration, 257
The Fortune at the Bottom of the Pyramid, 293*b*
Fossil fuels, 27, 122, 124, 127
France, 55, 147*b*, 300
French Revolution, 55

G

GAO. *See* General Accounting Office (GAO)
GDP. *See* Gross domestic product (GDP)
General Accounting Office (GAO), 344
Generic syllabus, sustainable engineering, 401–412
Genetically modified organisms (GMO), 18, 177
Geoengineering, 362, 363–366
  CCS, 363
  SRM, 364
Germany, 54, 237, 300, 303

Globalization, 51–62
Global Positioning System (GPS), 139
Global warming, 27, 35. *See also* Climate change
GMO. *See* Genetically modified organisms (GMO)
Golden Age, 42
Golden Rule, 329*b*–330*b*
GPS. *See* Global Positioning System (GPS)
Grasburg Mine, 386*b*
Great Chain of Being, 218*b*
Great Divergence, 26
Green chemistry, 245–263
  policy implications, 261–263
  principles of, 248*b*–249*b*
Greenhouse effect, 35, 36*f*. *See also* Climate change
Gross domestic product (GDP), 22, 23*f*
Gulf of Mexico, 40

H

Hall, Peter, 368
Hard power, 303, 304*b*–306*b*
Hart, Stuart, 293*b*
Haussmann, Baron, 327
Hazard, 133*b*
  moral, 364
Heine, Heinrich, 159*b*
Hierarchy of control, 183*b*
High modernism, 367–370
Hoplite war, 301
Human, as design space, 239–243
Human systems, nature shift to, 235–237, 236*f*
Hybrid systems, 227
Hydraulic engineering, 342, 353
Hydrologic civilizations, 342

I

IBM, 230*f*, 396, 398
  supercomputer simulation, 316
  virtual cat brain, 238
Ideologies, 110–111
India, 91, 132, 134
  chemical firm's incident in, 245, 247
  McDonald in, 339
Indian Ocean, 53
Industrial ecology, 189–212
  LCA, 198–209
  systems engineering, 210–212
  themes of, 192–194
Industrial materials, 122
Information and communication technology (ICT), 227, 229–239, 265–296
  as Anthropocene theme, 62, 65–67
  as bottom of the pyramid, 293*b*–294*b*
  changes in, 231–233
  culture and, 233
  firm-based perspective, 268*b*–270*b*
  framing, 267–272
  infrastructure, 272, 276–277
  nature shifts, 235–237, 236*f*
  services, 277–280
  as urban trend, 74–75
Information consumption, 62, 65–67. *See also* Information and communication technology (ICT)
Infrastructure operations, ICT, 272, 276–277
Innovation, 168–169
Integrity, and leadership, 393
International Standards Organization (ISO), 199
Internet-enabled personal computers, 172–173, 173*f*
Internet Protocol technology, 396
Interstate Highway System, 154
Iraq War, 306
ISO. *See* International Standards Organization (ISO)

J

Jacobs, Jane, 368
Japan, 57, 194
  gunpowder in, 185
  immigration in, 85
  incremental improvement processes and, 169
  manufacturing in, 176
  mobile telephone subscription in, 172, 172*f*
  robotics in, 227
  robot in, 320
  wireless infrastructure in, 174
Jeffersonian agrarianism, 156
Jencks, Charles, 368

K

Kondratiev waves, 144*b*, 156, 157*f*, 174, 176

L

Lake Okeechobee, 50, 51, 54
Landline telephone, 231
The Language of Post Modern Architecture, 368
LAR. *See* Lethal autonomous robots (LAR)
Leadership
  integrity and, 393
  learning and, 393–397
  overview, 391–393
  skills, required for, 398–400
Learning, and leadership, 393–397
Legacy system, 73
Leopold, Aldo, 337
Lethal autonomous robots (LAR), 319–323
  policy response matrix, 321*f*
Life cycle assessment (LCA), 198–209
  data, 203–204

goals and scope of, 199–201
impact analysis, 201–202
matrix structures, 205–207, 206*f*, 208*f*
methodologies, 203–204
results, 202
simplified, 198*f*
streamlining, 205–209
as a tool, 205
Life expectancy, 82, 83–84, 84*f*
Linear regime, 99
Locked in technology, 174, 372, 383–384
Love Canal, 245

M

Machine, brain and, 228*b*–229*b*
Macroethics, 326–340. *See also* Ethics
challenge of, 335–340
Manifest Destiny, 153–154
Marsh Arab societies, 348
Marxism, 56
Master Equation, 67–69
environmental, 68, 68*f*
sustainable engineering, 69, 69*f*
Materials. *See also* Resources
categories, 122
human mobilization of, 123, 123*f*
The Matrix, 218*b*
McDonald, 339
McNeill, J. R., 221
Megacity, 70, 72–73. *See also* Urbanization
Metals, 122
Metaphors, analogies and, 234*b*
Methane resources, 127
Military determinism, 143*b*
Military implications, of railroads, 154
Military operations, 299–323
cyborg insects, 310, 313–315
RCS, 308*f*, 310
RFID, 311
RMOC, 308*f*, 310
RMT, 307, 308*f*
RNC, 307, 308*f*, 309–310
telepathic helmets, 316–319
Mining, 385*b*–386*b*
Ministry of Industry and International Trade, in Japan, 194
Mobile telephone, 172, 172*f*, 173, 20–201
Models/modeling, 107–111
assumptions, 109
misuse of, 109–110
properly created, 108
reductionist approaches, 109
use of, 108
Mohamed, Mahathir bin, 374
Money, and technology, 144, 145*b*–147*b*
Monsanto, 63*b*–65*b*
Montreal Protocol, 387–388
Moore, Gordon E., 227
Moral hazard, 364
Moral philosophy, 327
Moses, Robert, 367
Multiculturism
and globalization, 51–62
and technological evolution, 57–59

N

Nanotechnology, 219–220, 221*f*
elements of, 220
stages of, 220
NaOH. *See* Sodium hydroxide (NaOH)
Napoleon III, 327
National Environmental Policy Plan, 195
National Environmental Policy Plan Plus, 195
National security, 299–323
cyborg insects, 310, 313–315
RCS, 308*f*, 310
RFID, 311
RMOC, 308*f*, 310
RMT, 307, 308*f*
RNC, 307, 308*f*, 309–310
telepathic helmets, 316–319
National Society of Professional Engineers Code of Ethics, 331*f*
Nation-state, 55
governance, changes in, 59–62
Natural disaster, 86
countries hit by, 87*f*
financial losses from, 86*f*
Nature, value system of, 114–117, 116*f*
Nature shifts, to human systems, 235–237, 236*f*
The Netherlands, 54
bicycle use in, 75
industrial ecology, 195
Network-centric firms, 288–296, 290*f*. *See also* Virtual office
collapse of boundary, 292, 295
conceptual model, 291
corporate efficiency, 291–292
time and, 292
Nitrate ions, in drinking water, 39–40
Nitrogen cycle, 39–40, 39*f*
for bio-based products, 44*f*
Nitrous oxide ($N_{16}O$), 39
Nokia, 398
Nongovernmental organizations (NGO), 60
Nonlinear regime, 99
Nuclear materials, 122
Nuclear weaponry, 301

O

Olmsted, Fredrick Law, 159*b*
Oppenheimer, Robert, 155, 301
Orgware, 16*b*
Our Common Future, 114, 115, 327

P

"Pay-as-you-go" offerings, 171
PCSD. *See* President's Council on Sustainable Development (PCSD)
Peace of Westphalia, 55
Pharmaceuticals, 256–261
Phytoplankton, 363
Plato, 234*b*
Plug-in automobiles, 104*b*–105*b*
Population
  age related issues, 78–85
  aging, 84–85
  life expectancy, 82, 83–84, 84*f*
  risk profile of, 85–87
  trends, 69–87
  urbanization and, 70–78
Prahalad, C. K., 293*b*
Precautionary Principle, 337
President's Council on Sustainable Development (PCSD), 119*b*–120*b*
Product design, 163–168
  DFA, 164–165
  DFD, 165
  DFE, 165
  DFM, 165–166
  DFR, 166
  DFRC, 166
  DFS, 166
  DFSM, 166–167
  environment and sustainability in, 168*f*
  gate process in, 167–168
Protestant Christianity, 54–55
Protestantism, 55

Q

Quads, 27

R

Radio frequency identification (RFID), 311
Railroads, 147–156
  capitalism and, 152–153
  communication and, 150
  concept of, 148
  cultural effect of, 155–156
  economic effects of, 153–154
  environmental impact of, 151–152
  industrial time and, 148–149
  landscape and, 152
  military implications of, 154
  psychological effect of, 154–155
  religious meaning of, 155
RCA, 396
Real time technology and policy assessment (RTTPA), 181, 183–184
Reductionist approaches, 109
Redundancy, designing, 384
Reflexivity, 240
Reformation, 54
Registry, 222–223
Reserves
  concept of, 124
  sustainability and, 122–128
  vs. resources, 125*b*–126*b*
Resiliency, designing, 386–387
Resource base, 126*b*
Resources
  concept of, 124
  vs. reserves, 125*b*–126*b*
Resource use, 68
Revolutions in Civilian Systems (RCS), 308*f*, 310
Revolutions in Military Operations and Culture (RMOC), 308*f*, 310
Revolutions in Military Technology (RMT), 307, 308*f*
Revolutions in Nature of Conflict (RNC), 307, 308*f*, 309–310
RFID. *See* Radio frequency identification (RFID)
Rights ethics, 332*f*, 333
Risk, 134*b*
RMOC. *See* Revolutions in Military Operations and Culture (RMOC)
RMT. *See* Revolutions in Military Technology (RMT)
RNC. *See* Revolutions in Nature of Conflict (RNC)
Robotics, 223, 226–227, 226*f*
Robots, 319–323
Roosevelt, Franklin D., 293*b*
Roosevelt, Teddy, 151
RTTPA. *See* Real time technology and policy assessment (RTTPA)
Russia, 60, 95
  DDOS attack of, 74, 309
  mobile telephone subscription in, 172, 172*f*

S

Sawgrass, 354
SCADA. *See* Supervisory control and data acquisition (SCADA) systems
Scale-free networks, 94, 95*f*
Scenario planning, 281*b*–282*b*
Services, ICT, 277–280
SETAC. *See* Society of Environmental Toxicology and Chemistry (SETAC)
Ship technology, 53
Shop floors, 182*b*–183*b*
Singer, Peter, 311
Skills, required for leadership, 398–400
Smart agriculture, 265
Smart grid, 127, 153, 176, 177, 265
Smith, Adam, 292, 323

Social equity, 117–118, 117*f*
Social ethics, 334
Social networking, 232, 232*f*, 266, 296, 365
  implications, 268
Society of Environmental Toxicology and Chemistry (SETAC), 199
Sodium hydroxide (NaOH), 363
Soft power, 303, 304*b*–306*b*
Software, 16*b*
Solar Radiation Management (SRM), 364
Southern Bell, 396
Southern Florida. *See* Everglades, Southern Florida; Florida
South Florida Ecosystem Restoration Task Force, 354
Soyuz spacecraft, 60
Spectrum, 240
SRM. *See* Solar Radiation Management (SRM)
Static complexity, 101, 101*f*
Sulfur aerosols, 364
Supervisory control and data acquisition (SCADA) systems, 287
Sustainability
  defined, 114
  egalitarianism and, 117–118, 117*f*
  environmental quality and, 115–117, 116*f*
  operational definition of, 118, 120
  as social myth, 121–122
  technology and, 67–69
Sustainable chemistry, 245
Sustainable development. *See also* Sustainability
  defined, 114
Sustainable engineer, 128–131
  in economic domain, 129
  in environmental domain, 129–130
  responsibility, 128
  in social and cultural domains, 130–131
Sustainable engineering, 61–62
  concept of, 16*b*–17*b*
  generic syllabus, 401–412
  macroethics of, 326–340
  urbanization and, 72–78. *See also* Urbanization
Synthetic biology, 222–223
Syr Dar'ya, 346
Systems
  in linear regime, 99
  in nonlinear regime, 99
  simple *vs.* complex, 92–100
  variation in, 92–93, 93*f*
Systems engineering, 210–212
  alternative solution, 211
  determining goals, 210
  initiating action, 212
  ranking alternative solution, 210–211

T

Tata Nano, 91, 132, 134
"Taylorism" industrial efficiency techniques, 158
TCS. *See* Triclosan (TCS)
"Technolithic," 141
Technological determinism, 143*b*–144*b*
Technological evolution
  and multiculturism, 57–59
  and sustainability, 67–69
Technological optimism, 144*b*
Technology, 138–160
  aspects of, 139*b*–141*b*
  growth of, 171
  human evolution and, 141, 142*f*
  implementation process, 169–170, 170*f*
  levels of, 178–184
  money and, 144, 145*b*–147*b*
  perception and, 159*b*–160*b*
  railroads as, 147–156
  research and development phase of, 168–169
  saturation point, 172, 174–176
  and sustainability, 67–69
Telegraph technology, and railroads, 150
Telepathic helmets, 238, 316–319
  policy response matrix, 317*f*
Terrorism, 58
Time, and railroads, 148–149
Time at desk (TAD) techniques, 284
Toxicity, 133*b*
Transhumanism, 216, 218*b*–219*b*
Triclocarban (TCC), 257–261, 258*f*
Triclosan (TCS), 257–261, 258*f*
Triple bottom line, ICT, 285–287. *See also* Virtual office
Twitter, 266, 315, 318. *See also* Social networking

U

United Nations, on urban population, 70–71, 71*f*–72*f*
United Nations Food and Agriculture Organization, 346
United States, age distribution in, 81, 81*f*
Uranium resources, 127
Urbanization, 70–72
  autonomic city, 76*b*
  bicycles, use of, 75, 77
  ICT and, 73–75
  mischaracterization, 73
  and sustainable engineering, 72–78
Utilitarianism, 332–333

V

Virtual office, 280, 282–296, 283*f*. *See also* Network-centric firms
corporate efficiency, 291–292
economic benefit of, 285–286
environmental implications, 287
firm level implications, 288–296
quality-of-life, 286
social benefit of, 286–287
time and, 292
work and family, integration of, 292, 295
Virtue ethics, 332*f*, 333

W

Walmart, 291–292
War porn, 323
Water, 31–35
embedded. *See* Embedded water
global freshwater use over the past 314 years, 31*f*
Water Resources Development Act, 354
WCED. *See* World Commission on Environment and Development (WCED)
Wetware, 16*b*
Wicked complexity, 102–103
Wired for War, 311
Wireline communication, 231
Wood creosote, 152
World Business Council on Sustainable Development, 385*b*
World Charter for Nature, 337
World Commission on Environment and Development (WCED), 114
World Commission on Sustainable Development Report, 327
Worldviews, 110–111
Writing, 46*b*–48*b*

X

X-ware, 16*b*