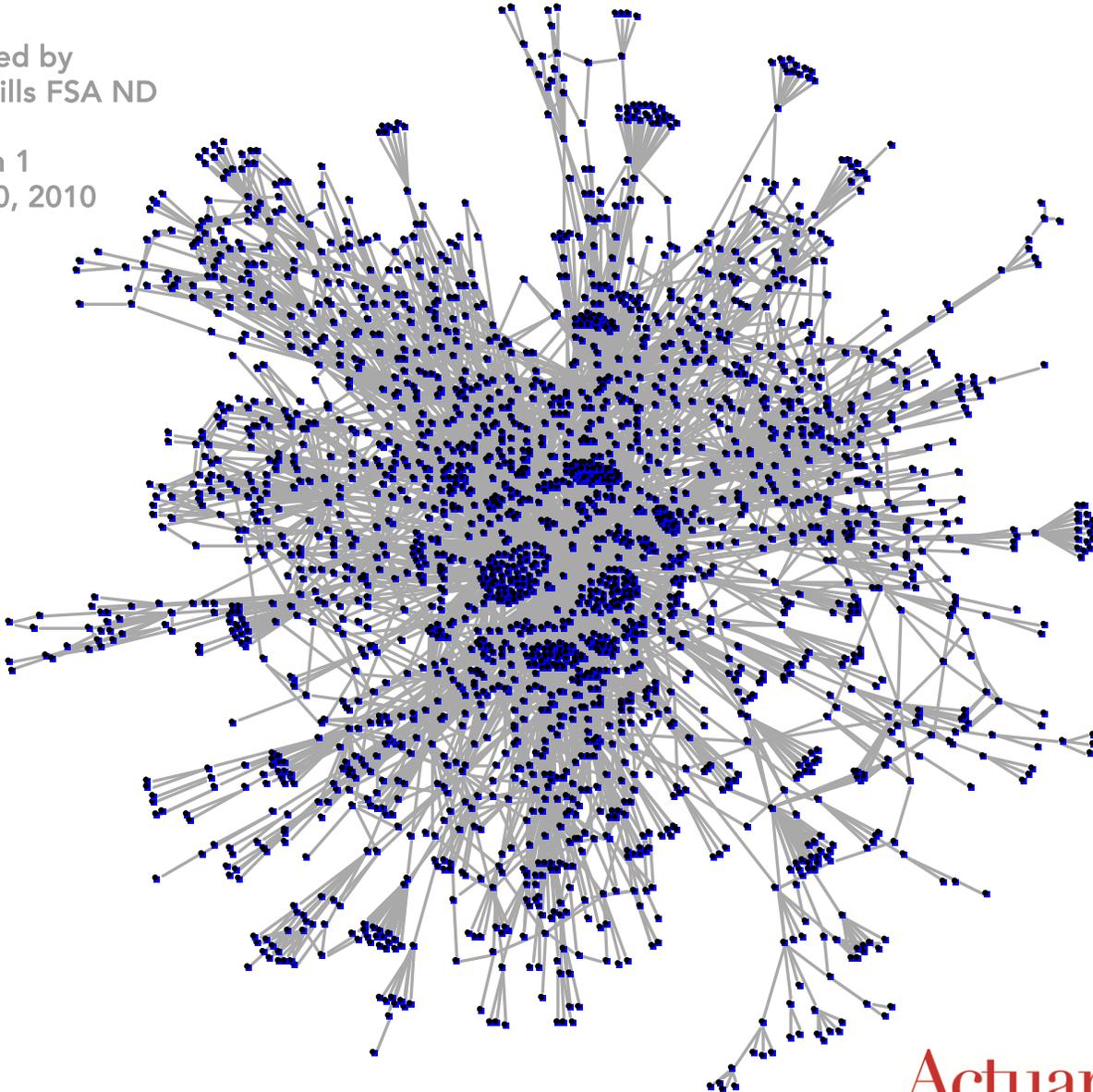


Complexity Science

An introduction (and invitation) for actuaries

Prepared by
Alan Mills FSA ND

Version 1
June 10, 2010



Actuaries
Risk is Opportunity.®

The opinions expressed and conclusions reached by the author are his own and do not represent any official position or opinion of the Society of Actuaries or its members. The Society of Actuaries makes no representation or warranty to the accuracy of the information.

The greatest challenge today, not just in cell biology and ecology, but in all of science, is the accurate and complete description of complex systems.

Edward O. Wilson, 1990

The central task of theoretical physics in our time is no longer to write down the ultimate equations but rather to catalogue and understand emergent behavior in its many guises, including potentially, life itself. We call this physics of the next century the study of complex adaptive matter ... We are now witnessing a transition from the science of the past, so intimately linked to reductionism, to the study of complex adaptive matter ... with its hope for providing a jumping-off point for new discoveries, new concepts, and new wisdom.

Robert Laughlin and David Pines, 2000

I think the next century will be the century of complexity.

Stephen Hawking, 2000

CONTENTS

	Acknowledgments	
	Introduction	1
PART I	A NEW SCIENCE	
ONE	Complexity Science	6
	Science of the real world	
TWO	Agent-based models	28
	The heart of Complexity Science	
PART II	COMPLEXITY SCIENCE MODELS	
THREE	Four archetypal models	49
	The four archetypal models of Complexity Science	
FOUR	Networks	53
	Agents and relationships	
FIVE	Cellular automata	77
	Agents, relationships, and behavior rules	
SIX	Artificial societies	119
	Agents, relationships, behavior rules, and an environment	
SEVEN	Serious games	156
	Agents, relationships, behavior rules, an environment, and a player	
PART III	AN INVITATION	
EIGHT	The complex systems actuary	172
	Vision for a new type of actuary	
NINE	Next steps	178
	Steps to develop complex systems actuaries	
	Top ten Complexity Science books	
	Essential resources	
	Finding the essential resources	
	Glossary	

ACKNOWLEDGMENTS

I am grateful to many for their encouragement, advice, and support:

- During my Complexity Science presentation at the 2009 Health Spring Meeting, **Judy Strachan** recognized how important this subject could be for actuaries, and afterwards encouraged me to write about it for the Society of Actuaries (SOA). Judy helped obtain funding for this report, and chaired the report's Project Oversight Group.
- After considering Judy's invitation, I initially doubted that many actuaries would see the potential of Complexity Science, and so initially declined to write about it for actuaries. **Steve Siegel** persuaded me to reconsider. He argued convincingly that many actuaries would welcome the insights and tools of Complexity Science to help them solve the ever-more complex problems they face.
- During the report's preparation, members of its Project Oversight Group gave consistently wise counsel, engaged support, and helpful critique. The members are:
 - **Judy Strachan**, Chair
 - **Steve Conwill**
 - **Syed Mehmud**
 - **Bernie Rabinowitz**
 - **Lee Smith**
 - **John Stark**
 - **Steve Siegel**, SOA Research Actuary
 - **Sara Teppema**, SOA Staff Fellow for Health
 - **Barbara Scott**, SOA Research Administrator
- **Dave Snell** – a fellow Complexity Science aficionado, experienced editor, and actuary – offered sound suggestions for improvement.
- **Scott Page**, a prominent Complexity Science expert, provided many detailed and helpful recommendations to improve the report.
- **Randall Rupper**, a full-time physician and part-time Complexity Science researcher, suggested several improvements.
- **Neil Cattle, Tom Conway, Rick Gorvett, Don Mango, Jon Palin, and Michael Shumrak** are actuaries who are Complexity Science pioneers. They shared their related experiences and contacts, and enthusiastically supported this report.
- **Lisa Canar**, my wife, read every word of the report and every line of computer code – many times – and offered helpful suggestions each time. Even more difficult to achieve, she and our small children, **Calun** and **Ariana**, stayed consistently supportive during the many evenings and nights of my absence while I wrote this report.

To all of you, my warmest thanks.

INTRODUCTION

A. MORE THAN AN INTRODUCTION

Complexity Science is a new way to grasp and manage reality. Not the simple reality of planetary motion and gambling dice that has been the study of traditional science. But the complex reality in which we live: the world of hurricanes and earthquakes, social reforms and economic upheavals, interest rate fluctuations, business cycles, and healthcare expenditure trends. In this report, I will introduce you to this surprising and useful new science.

But this report is more than an introduction. In it, I will also make an argument, and extend an invitation to you.

I will argue that today's complex world demands a new breed of actuary, a professional I call a 'complex systems actuary', who:

- **Understands the complex nature of social systems.** The social systems in which actuaries work (the worlds of insurance, pensions, investments, and health care) are not well-behaved like planets and dice. Rather, they are complex and unpredictable.
- **Applies new methods.** Traditional actuarial methods alone cannot model complex social systems. To grasp and manage the systems in which we work, we must augment our tools with the new methods of Complexity Science.
- **Expands the role of actuaries.** Actuaries need not be merely administrators within society's problem-prone systems. Rather, we also can use our unique knowledge and skill, augmented with insights and tools from Complexity Science, to help solve society's great problems and improve our social systems.

I will invite you to become such an actuary. No matter if you are a veteran or just starting out, you can learn to use the insights and tools of Complexity Science to address the most pressing problems of your employer, your clients, and society as a whole. Your fitness for our complex world may depend on whether you accept this invitation, or simply send regrets.

My affair with complexity

It was the summer of 2005 when, finally, I saw a gleam of hope that the US healthcare system can be understood. That summer I worked with the Center for the Study of Complex Systems at the University of Michigan in Ann Arbor, and I read Joshua Epstein and Robert Axtell's potent little book *Growing artificial societies*.¹

At least ten years before, I knew the US healthcare system was in trouble, but I could not understand why. As a healthcare consumer, actuary, and physician, I had experienced many sides of the healthcare problem: the lack of coverage, runaway expenditures, and tortuous reimbursement procedures.

After studying the problem and divers analyses of it, I came to a shocking realization: even though nearly everyone offered fixes for the healthcare system, *no one understood it*.

I assumed that the system was too complex to be understood, and that healthcare policy would forever be gropings of the blind. But then, in 2003, I read Stephen Wolfram's newly published book *A new kind of science*.² It was unlike anything I had seen, a brilliant light that led me to Ann Arbor to learn about Complexity Science. There, while reading Epstein and Axtell's book, I saw that the healthcare system can be modeled and understood, that clear-sighted healthcare policy is possible.

Since then, I have also learned that *all* the complex social systems in which actuaries work can be modeled and understood, but that for this purpose traditional actuarial methods are insufficient.

In this report, I will share with you what I have learned.

¹ Epstein & Axtell (1996), one of this report's *Top ten Complexity Science books*.

² Wolfram (2002), also one of this report's *Top ten Complexity Science books*.

B. MAJOR THEMES

Throughout this report I interweave five themes:

1. **Social systems are complex systems.** The social systems in which actuaries work are complex systems, with mechanisms dramatically different from those of simple systems such as planets and dice. To understand and manage the behaviors of such systems – this is society’s greatest challenge.
2. **We must study complex system behavior from the bottom up.** The behavior of a complex system arises from the bottom-up, from its components, the relationships among its components, and the behavior rules that the components follow. To understand and manage such systems, we must model them from the bottom up, using special methods of Complexity Science, rather than top-down traditional actuarial methods.
3. **Long-term prediction of complex systems is impossible.** The long-term behavior of complex systems – such as the fluctuations of financial market prices and health care trend rates – cannot be accurately predicted for more than short periods. Actuaries pursuing long-term prediction of complex systems are wasting time.
4. **Understanding and effectively managing complex systems is possible.** Though the long-term of complex systems cannot be predicted, their behavior can be understood and managed, like a farmer manages the cultivation of crops.
5. **Actuaries can help solve society’s great problems.** Using our unique skills and knowledge – along with the tools and insights of Complexity Science – actuaries can effectively address the great problems of complex social systems, and lead the development of new social policy, rather than merely administer existing problematic systems.

In exploring these themes, I will use many examples and issues from health care, because that is where my primary interest lies, and because the Society of Actuaries (SOA) Health Section funded this report. But I will also include examples from other actuarial areas, including pensions, investment, property-casualty, and reinsurance. The basic concepts, insights, and tools of Complexity Science apply to all complex systems and all actuaries.

Complexity science – an introduction (and invitation) for actuaries

C. ORGANIZATION

The report is organized in three parts and nine chapters:

- **Part I – A new science** (two chapters) gives an overview of Complexity Science and agent-based models, and shows how they relate to the problems of complex social systems.
- **Part II – Complexity Science models** (five chapters) introduces Complexity Science models and methods. It presents four archetypal models of Complexity Science, and provides the background and tools for you to start applying them in your work.
- **Part III – An invitation** (two chapters) presents a vision of a new type of actuary, the ‘complex systems actuary’, and proposes a plan to develop this professional.

Each chapter includes exercises to help you better understand its material. Answers to these exercises are in the document titled *Answers to exercises*, found on the SOA web page for this report.

D. SUPPLEMENTAL MATERIAL

Supporting the report is the following supplemental material:

- **Top ten Complexity Science books.** Among the scores of books about Complexity Science, ten stand out for actuaries. They are listed and annotated in the section titled *Top ten Complexity Science books* at the end of the report.
- **Essential resources.** To prepare this report, my first step was to search for material about Complexity Science that is relevant for actuaries. In this search, I reviewed thousands of potentially relevant journal articles, books, web sites, unpublished reports, and videos. I distilled the results of this search into about one hundred resources with brief annotations, found in the section titled *Essential resources* at the end of the report. The report’s bibliographic citations (found in footnotes) refer to these resources.

Because you may want to learn how to perform such a search, another section at the end of the report, titled *Finding the essential resources*, details how I performed the search, and a document on the SOA web page for this report, titled *Literature search results*, details the search results.

D. SUPPLEMENTAL MATERIAL CONTINUED

- **Computer code.** Until you use it, Complexity Science is useless. Perhaps the most important material to help you start using Complexity Science are the models and computer code supporting the examples in Part II of the report. They are found on the SOA web page for this report. I carefully documented the computer code to make it easier for you to apply Complexity Science models in your work.

To help you set up the computer platforms for the Complexity Science models, there is a document titled *Getting started with modeling platforms* on the SOA web page for this report.

- **Glossary.** The report includes a glossary of technical terms. When a technical term is first introduced, it is bold and in quotation marks, like ‘**new term**’.

E. LEARNING OBJECTIVES

After you read this report and work with the Complexity Science models, you will be able to:

- Explain the basic concepts, tools, insights, and results of Complexity Science.
- Discuss the most important resources, organizations, and people in Complexity Science.
- Understand and assess any book, journal article, or other material about Complexity Science.
- Develop models of complex systems using the modeling platforms Excel, igraph, and Repast Symphony.
- Explain why Complexity Science is important in actuarial work.
- Explain how an actuary can become a complex systems actuary, and why you might want to become one.

I hope you enjoy the report. Even more, I hope you become a complex systems actuary. If you start along this path, please let me know. (And please let me know if you have suggestions to improve this report.)

Alan Mills (alan.mills@earthlink.net)

June 10, 2010

PART I: A NEW SCIENCE

I am convinced that the societies that master the new sciences of complexity and can convert that knowledge into new products and forms of social organization will become the cultural, economic, and military superpowers of the next century.

Heinz Pagels. 1988¹

¹ Pagels (1988), page 53.

CHAPTER ONE: COMPLEXITY SCIENCE

The movement's nerve center is a think tank known as the Santa Fe Institute ... The researchers who gather there are an eclectic bunch, ranging from pony-tailed graduate students to Nobel laureates ... But they all share the vision of an underlying unity, a common theoretical framework for complexity that would illuminate nature and humankind alike. ... They believe that their application of these ideas is allowing them to understand the spontaneous, self-organizing dynamics of the world in a way that no one ever has before – with the potential for immense impact on the conduct of economics, business, and even politics. ... They believe they are creating, in the words of Santa Fe Institute founder George Cowan, 'the sciences of the twenty-first century.'
Mitchell Waldrop¹

A. INTRODUCTION

How is it that from our cataclysmic origin (the so-called Big Bang) only about 10^{17} seconds ago, we now have insurance companies, pension plans, health systems, and the Society of Actuaries, not to mention the New York Stock Exchange, the United Nations, and myriad other social systems? All this in the face of the immutable Second Law of thermodynamics that says we should instead be speeding toward a bland soup of hadrons and quarks.

Now that we have such social systems, how will they continue to develop? Will they smoothly progress toward ever-more cohesion, efficiency, usefulness, and complexity? Or, like 99 percent of all past life forms, will they soon go extinct?

Why do our social systems take the particular forms they do? Why do they go through sudden dramatic upheaval, such as the collapse of the Soviet Union, the 1987 stock market crash, or the recent financial meltdown? Can we predict the future of our social systems? Can we manage their evolution? Can we, should we, optimize their efficiency?

Such questions are the focus of '**Complexity Science**', a new field that studies universal principles common to all complex systems – business firms, countries, galaxies, roses, and children: how they form, how they evolve, how they die.

Though young, Complexity Science has much to say about such questions – and much that is relevant to actuarial work. This chapter presents an overview of Complexity Science, including its development, hallmarks, concepts, tools, and key results. The chapter ends with a review of current issues in Complexity Science and exercises to sharpen your understanding of this new field.

¹ Waldrop (1992), pages 12-13. This book is one of this report's *Top ten Complexity Science books*.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

B. BIRTH OF A NEW SCIENCE

It was 1982, in a cafeteria at Los Alamos National Laboratory (LANL), during the weekly luncheon of the lab's Senior Fellows. George Cowan voiced a radical idea that would give birth to a new science.

The Senior Fellows were semi-retired LANL senior scientists who nevertheless were current with the latest scientific concepts, such as chaos and dynamical systems. George Cowan, one of them, had worked at LANL for 39 years and had managed its research (see sidebar). He proposed that the group establish a new trans-disciplinary research institute, an institute devoted to the study of real-world complex systems in nature and society; that for such study the institute would assemble all types of scientists – physicists, mathematicians, economists, biologists, cognitive psychologists, and social scientists; and that to enable such study the institute would build an advanced computer facility. The Senior Fellows greeted Cowan's idea with resounding enthusiasm and support.

In February 1987, the Santa Fe Institute (SFI) opened. Located in Santa Fe, about 35 miles from LANL, SFI was first housed in an old adobe structure, a charming place that was formerly a nunnery, with ceilings held up by rough-hewn fir beams, small bedrooms that were converted into offices, and a former chapel with stained-glass windows that became the conference room – a fit setting of stability, calm, and tradition to contrast with SFI's paradigm-breaking ambitions. Cowan was its president, and its board chair was legendary Nobel laureate Murray Gell-Mann.

Today, more than two decades later, SFI has become the vibrant research center Cowan imagined. It is in a larger building, with 35 year-round resident interdisciplinary researchers, 70 researchers in summer, 60 external faculty members, and 25 administrative staff members. Each year, it hosts two dozen workshops, and an annual summer school with 150 students. It researches the range of complex systems, from cells and biological systems to business firms and economic systems.

Through its Business Network, SFI works with business organizations to make sense of complexities they face. Deloitte Touche Tohmatsu and Towers Watson are members.³

A determined man

Like most people associated with Complexity Science, George Cowan is multi-faceted. As one of the original Manhattan Project researchers and head of research at LANL, he created deadly weapons; yet, he also founded the Los Alamos National Bank, and was the initial inspiration for the Santa Fe Institute.

In his book *Complexity*, Mitchell Waldrop wrote about Cowan:

“Cowan was the one who had conceived the institute in the first place. He was the one who had envisioned a science of complexity before anyone had even known what to call it. He was the one who had done more than anyone else to make the Santa Fe Institute happen, to make it the most intellectually exciting place that any of them had ever been in.

... he was a retiring, soft-spoken man who managed to look a bit like Mother Teresa in a golf shirt and unbuttoned sweater. He was not noted for his charisma; in any given group he was usually the fellow standing off to one side, listening. And he was certainly not known for his soaring rhetoric. Anyone who asked him why he had organized the institute was liable to get a precise, high-minded discussion of the shape of science in the twenty-first century and the need to take hold of scientific opportunities ...

Only slowly, in fact, would it begin to dawn on the listener that Cowan, in his own cerebral way, was a fervent and determined man indeed. He didn't see the Santa Fe Institute as a paradox at all. He saw it as embodying a purpose far more important than George A. Cowan, Los Alamos, or any of the other accidents of its creation ... To Cowan, it was a chance for science as a whole to achieve a kind of redemption and rebirth.”²

² Waldrop (1992), pages 54-55 and 336. This book is one of this report's *Top ten Complexity Science books*.

³ To learn more about the SFI Business Network, visit “www.santafe.edu/network”.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

B. BIRTH OF A NEW SCIENCE CONTINUED

In addition to SFI, there are now many active Complexity Science research groups, including:

- Center on Social and Economic Dynamics (Brookings Institution)
- New England Complex Systems Institute
- Center for the Study of Complex Systems (University of Michigan)
- Center for Social Complexity (George Mason University)
- Institute on Complex Systems (Northwestern University)⁵

For the first time in history, these groups are focused squarely on complex systems, a subject that traditional science has avoided: Physics, with its emphasis on simple systems, has historically defined itself to avoid complexity. Biology has concentrated on specific observation, with little theoretical discussion of general phenomena such as complexity. With ‘general systems theory’ in the 1960s, social science briefly attempted to examine complexity in human organizations, but its results were ineffective. In the 1970s, the new areas of fractals (see Chapter six) and ‘**chaos theory**’ (see sidebar) addressed one type of complexity, but they focused only on features of complex systems that could be summarized in mathematical equations.⁶

C. COMPLEX SYSTEMS

You may wonder, “What is a complex system?” That is a hard question, one that has even Complexity Science experts baffled. For example, the experts John Miller and Scott Page write, “Rather than venturing further on the well-trodden but largely untracked morass that attempts to define complex systems, for the moment we will rely on Supreme Court Justice Stewart’s words ... on a case dealing with obscenity: ... ‘I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it.’”⁷

However, to give you some idea of the subject of complex systems, let’s step briefly into the morass:

Chaos theory

One day in 1961, the meteorologist Edward Lorenz was using a computer to model weather with a set of twelve equations. He had run the equations to evolve a weather pattern for a particular time period, but wanted to see the end of the pattern again. To save time, rather than start the simulation at the beginning of the time period, he started it midway through, using parameters from his previous printout for the starting point.

An hour later, when he returned, he was surprised: rather than reproduce the prior results, the new results ended up wildly different. He found the reason: the parameters he entered for the second run were accurate only to three decimal places, whereas the initial run had results accurate to six decimal places. That small difference in initial conditions led to wildly different weather.

This effect, sensitive dependence on initial conditions, became known as the butterfly effect: a butterfly’s flight in Texas today can produce a tiny atmospheric change that a month later produces a storm in Thailand.

This result led to the conclusion that it is impossible to predict the weather for more than a few days, and to chaos theory.

Chaos theory is the mathematical study of dynamic systems that are highly sensitive to initial conditions (‘**chaotic systems**’). Many complex systems – such as the weather, populations, neurons, and economic systems – can also be chaotic systems.⁴

⁴ To learn more about the origins of chaos theory, see Gleick (2008).

⁵ For information about these and many more centers, see Sanders & McCage (2003).

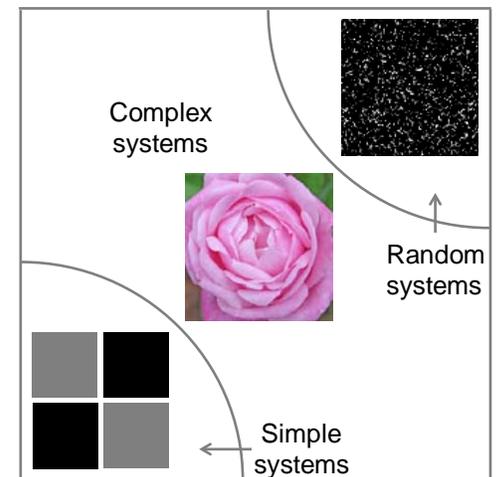
⁶ Wolfram (2002), pages 861-863. This book is one of this report’s *Top ten Complexity Science books*.

⁷ Miller & Page (2007), page 3, one of this report’s *Top ten Complexity Science books*.

C. COMPLEX SYSTEMS CONTINUED

Let's start by defining basic terms:⁸

- **'set'**: *A collection of objects.* For example, the moon, a cat, and a pixel are objects that we can call a set. The objects need not be related to one another.
- **'system'**: *A set whose objects are related to one another.* For example, the grid of pixels on your computer monitor naturally form a system: They are related objects because they are part of one lattice network, working together to perform a particular function.
- **'dynamic system'**: *A system together with a behavior rule that causes the state (ie, an attribute) of at least one of its objects to change over time.* Because a pixel can assume a variety of intensities and colors, and because your computer can send a behavior rule to pixels causing them to change their state, the system of monitor pixels together with the behavior rule is a dynamic system. Note that to define a dynamic system, we introduced the concept of an object's state, and the concept of a rule that governs the object's behavior over time.
- **'simple system'**: *A dynamic system for which the state changes of its objects are relatively uninteresting.* For example, if all your monitor pixels were to simultaneously flash black then gray then black then gray, one flash per second non-stop, they, together with their behavior rule, would be relatively boring and thus constitute a simple system. To define a simple system, we have now introduced human judgment about what is interesting.
- **'random system'**: *A dynamic system for which the state changes of its objects appear to be random.* For example, if your monitor pixels were to continuously produce 'snow', covering your monitor with random pattern-less white dots, then they, together with the 'snow' behavior rule, would be a random system. Again, such a system is relatively uninteresting.



Now suppose a brilliant budding flower appears on your monitor, opening magically into a glorious rose. This dynamic system – the collection of related objects (pixels) and their new behavior rule – captures your interest. It has become a **'complex system'**: *an interesting dynamic system.*

⁸ I do not mean for these definitions to have logical rigor; rather, I merely want to develop a helpful sketch of 'complex systems'.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

C. COMPLEX SYSTEMS CONTINUED

Now imagine your entire visual field as a vast monitor. Some things you see are simple and uninteresting, some appear random and are also uninteresting, but some – the pattern of a snowflake, the swirl of a galaxy, or the development of a child – are absorbingly interesting. These are complex systems. No longer are their objects simple pixels fixed in a grid; rather, they are molecules, stars, and biological cells that take myriad states and move. Other complex systems you will see are financial markets with dramatically fluctuating prices, companies that come and go, health care systems with unpredictable expenditure trends, and governments that rise and fall.

Thus, complex systems are collections of related objects with intriguing patterns of evolution that we find somewhere between randomness and simplicity. In this report, we call the objects of complex systems ‘agents’ (more about them in the next chapter); and to the agents, their relationships, and their behavior rules, we add one more ingredient, an ‘environment’, in which the agents move and with which they interact. Agents, relationships, behavior rules, and an environment: these are the basic elements of all complex systems.

Much of Complexity Science deals with a special subset of complex systems found particularly in social settings, ‘**complex adaptive systems**’. Such complex systems change their behavior to adapt to changes in their environment. Physician groups are an example of a complex adaptive system: they commonly change their diagnostic and prescribing behavior in response to changing insurance reimbursement policies.

I hope this sketch has given you a sense for complex systems, a concept that eludes definition. Perhaps it is the human element, our subjective evaluation of a dynamic system as interesting or not, that makes an objective definition so elusive. Just as the movement of planets in our solar system was at one time mysterious and intriguing, what counts as a complex system today may become tomorrow’s simple system. Similarly, what today appears random may tomorrow be found to have interesting patterns of a complex system (see the sidebar).

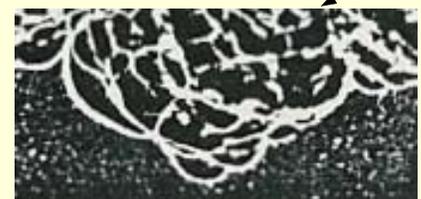
Randomness redux

For centuries, scientists have vigorously debated the structure of high-speed turbulence. Some held that even at high speed, turbulent fluids consist of little waves and mini-vortices, as in Leonardo da Vinci’s drawing of turbulence.



Others were emphatic that this was impossible: there is no structure, there is randomness and nothing more. This debate led physicist and Nobel laureate Richard Feynman to say, “Turbulence is the greatest puzzle in classical physics.”

In 1970, Gary Brown and Anatol Roshko solved the puzzle. They took pictures of two gases, helium and nitrogen, flowing next to each other at high speed, forming turbulence. Their famous images show that, rather than being completely random, turbulence consists of repeating nested swirling patterns, especially at its edges. Thus, what was once thought a purely random system is now understood to be a complex system. Da Vinci was right.⁹

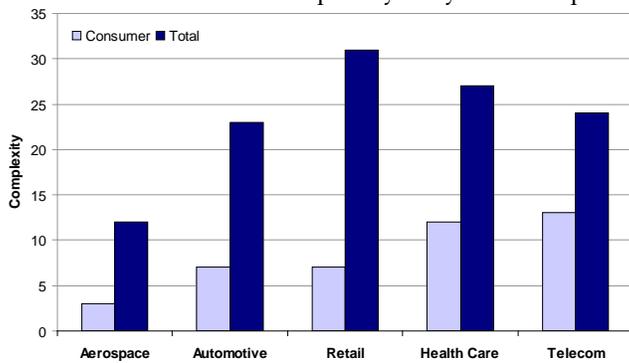


⁹ The drawing by Leonardo da Vinci is ‘turbolenza’ found in his notebook called the *Codex Leicester* (now owned by Bill Gates). The images of turbulence are in Brown and Roshko’s paper titled, “On density effects and large structures in turbulent mixing layers”, published 1974 in the *Journal of Fluid Mechanics*. See Bass (1999), pages 92-95 for more about the story, and how it relates to randomness in financial markets.

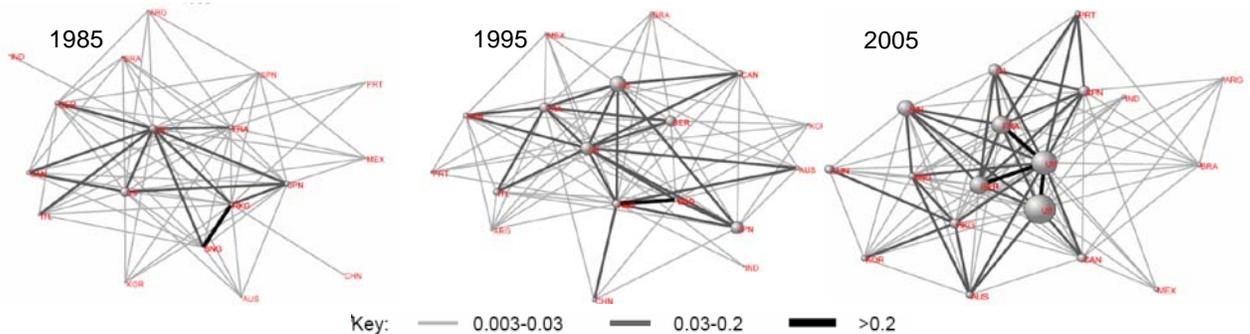
D. COMPLEXITY

Even though we cannot accurately define a complex system, we can measure its complexity – ironically, in perhaps too many ways (see sidebar).

For example, using a ‘transaction information’ complexity measure, the chart below compares the complexity of five US service-oriented complex systems. The measurements are taken from two perspectives: the information required by the system as a whole, and the information required by the system’s consumers. For example, from the perspective of the healthcare system as a whole, to determine the elements of an average medical expenditure transaction, one must answer about a billion binary questions.¹¹ As the chart shows, by this measure of complexity, the retail system is the most complex, but its complexity is mostly hidden from consumers. Not so for health care: it has one of the highest ratios of consumer complexity to system complexity.



The chart below shows the evolution of the international financial network’s complexity, from 1985 to 2005. The network nodes (vertices) are countries; node size is proportional to total external financial stocks held by each country; and the thickness of links between nodes is proportional to the ratios of external bilateral financial stocks to GDP. Clearly, complexity has increased.¹²



¹⁰ See Mitchell (2009), pages 94-111, and Érdi (2008), pages 201-214.

¹¹ Basole & Rouse (2008).

¹² Haldane (2009)

Measures of complexity

There are at least forty measures of a complex system’s complexity. These measures relate to how hard it is to describe the system, how hard the system is to create, or the degree of its organization.

Following are examples:¹⁰

- **Transaction information:** The number of bits of information required to identify the elements of a typical system transaction.
- **Network complexity:** There are many measures of network complexity that we will explore in Chapter four. One is the average number of connections per network vertex.
- **Degree of hierarchy:** The nestedness, or levels of hierarchy, within a system. More complex systems have more levels.
- **Algorithmic information content:** The number of bits in the shortest computer program that completely describes the system.
- **Logical depth:** The number of steps a Turing machine would take to construct the series of 0s and 1s that completely describes a system. This is a measure of how difficult it is to construct a system.
- **Statistical complexity:** The minimum amount of information about a system’s past behavior required to predict its near-term future statistical behavior.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

E. HALLMARKS

In addition to a focus on complex systems, three other hallmarks differentiate Complexity Science from traditional science:

- **Trans-disciplinary.** Recognizing deep similarities among complex systems throughout nature and society, Complexity Science no longer divides our study of reality into the areas of traditional science, such as physics, biology, and social science. Rather, it cultivates a new breed of scientist who can explore complex systems in any setting.
- **Constructive.** Complexity Science models complex systems from the bottom up: starting from the interactions of individual objects following their behavior rules, it constructs complex systems that usually have characteristics not predictable from an analytic study of their components.

This approach is in stark contrast to most sciences, where either a purely top-down aggregate statistical perspective is favored (as in the social sciences and with actuarial models) or where a reductionist approach predominates – tearing systems down to their components, but not building them back again (as in physics and biology). In the reductionist approach, there is an underlying assumption that if one understands a system’s components, then one understands the system as a whole, an assumption Complexity Science has shown to be false (see sidebar).

- **Computer-based.** Because most complex systems consist of thousands or millions of interacting objects, following a variety of behavior rules, computers are necessary to model their interactions and to construct the resulting system behavior as a whole. Traditional mathematics is insufficient for such a task. Indeed, computer availability was a major reason Complexity Science arose.

Writing about the convergence of scientific disciplines in Complexity Science, Heinz Pagels wrote, “The material force behind this change is the computer, the instrument of the sciences of complexity. ... The computer as a research instrument provides us with a new way of seeing reality, and the architectonic of the sciences must change accordingly.”¹⁴

Complete mystery

“In the existing sciences much of the emphasis over the past century or so has been on breaking systems down to find their underlying parts, then trying to analyze these parts in as much detail as possible. And particularly in physics this approach has been sufficiently successful that the basic components of everyday systems are by now completely known. But just how these components act together to produce even some of the most obvious features of the overall behavior we see has in the past remained an almost complete mystery.”

Stephen Wolfram¹³

¹³ Wolfram (2002), page 3.

¹⁴ Pagels (1988)

F. COMPARISON WITH OTHER FIELDS

Complexity Science is often confounded with other fields, even though it is fundamentally different. Following are brief comparisons. My hope is that from seeing what Complexity Science is not, you will better understand what it is.

- **‘Artificial intelligence’**: This field’s goal is to develop computers that think (‘strong AI’), or that perform specific reasoning tasks (‘weak AI’). Even though Complexity Science studies the brain as one complex system, it also studies many others. And its goal is understanding, rather than performance.
- **‘Artificial life’**: This field shows that computer programs can emulate certain features of living organisms. Again, although Complexity Science studies living organisms as complex systems, and employs computer programs to model them, it has a broader scope.
- **‘Catastrophe theory’**: Popular in the 1970s, this field is a branch of mathematics that studies how large discrete changes (catastrophes) can appear in solutions of continuous equations with only small parameter changes. Complexity Science, by contrast, is not a branch of mathematics, and indeed views traditional mathematics as only one tool – usually an inadequate one (see sidebar) – to study complex systems.
- **‘Chaos theory’**: Chaos theory is a branch of mathematics that studies the mathematical characteristics of dynamic systems that are highly sensitive to initial conditions. Although many complex systems are sensitive to initial conditions, Complexity Science’s interest in them goes beyond their mathematical characteristics.
- **‘Computational complexity theory’**: A branch of computer science, this field classifies computational tasks according to their inherent difficulty. In contrast, Complexity Science is not a branch of computer science. Neither is it focused on the relative difficulty of computer programs.
- **‘Cybernetics’** and **‘Systems dynamics’**: Originating in electrical engineering, these related fields deal with the aggregate non-linear behavior of systems characterized by feedback loops. Their perspective is top-down, looking at a system’s aggregate behavior, whereas Complexity Science studies systems from the bottom up, how their behavior arises from their components and the behavior rules the components follow.

Limitations of mathematics

“Three centuries ago science was transformed by the dramatic new idea that rules based on mathematical equations could be used to describe the natural world. ...

If theoretical science is to be possible at all, then at some level the systems it studies must follow definite rules. Yet in the past throughout the exact sciences it has usually been assumed that these rules must be ones based on traditional mathematics. But the crucial realization ... is that there is in fact no reason to think that systems like those we see in nature should follow only such traditional mathematical rules. ...

One might have thought that with all their successes over the past few centuries the existing sciences would long ago have managed to address the issue of complexity. But in fact they have not. And indeed for the most part they have specifically defined their scope in order to avoid direct contact with it. For while their basic idea of describing behavior in terms of mathematical equations works well in cases like planetary motion where the behavior is fairly simple, it almost inevitably fails whenever the behavior is more complex.”

Stephen Wolfram¹⁵

¹⁵ Wolfram (2002), pages 1 and 3.

F. COMPARISON WITH OTHER FIELDS CONTINUED

- **‘Dynamical systems theory’**: A branch of applied mathematics, this field studies systems that can be modeled with a particular class of mathematical equations (differential equations or difference equations). Again, Complexity Science is not a branch of mathematics, and the behavior of complex systems that it studies generally cannot be captured by such mathematical equations.
- **‘Experimental mathematics’**: A branch of mathematics, this field uses computers and numerical computation to investigate mathematical objects and properties. Generally, the field studies objects and properties that have already been investigated using traditional mathematics. By contrast, Complexity Science employs computers to investigate properties of complex systems that are usually inaccessible to traditional mathematics.
- **‘Fractal geometry’**: This field studies shapes in nature, and shows that many are not regular or smooth, but rather are nested shapes with intricate patterns. The study of Complexity Science goes far beyond shapes found in nature. Using simple rules, Complexity Science can generate the nested patterns found in natural systems, but also can produce many other patterns.
- **‘Game theory’**: A branch of applied mathematics, this field studies behavior in strategic situations where one person’s (or organization’s) choices depend on the choices of others. Complexity Science may employ game theory results in developing its behavior rules (see Chapter five, section B)
- **‘General systems theory’**: Popular in the 1960s, this field studies the general principles of social system functioning. Although the field may be considered a precursor to Complexity Science, its practitioners did not succeed in convincing others of its practical value.
- **‘Non-linear dynamics’**: A branch of mathematics, this field analyzes non-linear mathematical equations. Complexity Science is not a branch of mathematics, and the behavior of the complex systems it studies often cannot be captured by such mathematical equations.¹⁶

¹⁶ For a more detailed discussion of differences between Complexity Science and other fields, see Wolfram (2002), pages 12-16.

G. BASIC CONCEPTS

This section describes seven basic concepts in Complexity Science that you will encounter often.

1. Computation

One may think of the aggregate behavior of a complex system as a ‘**computation**’, in which each of its agents carries out (or computes) its behavior rule, the way a computer carries out its program. Equivalently, we can say that the complex system is processing ‘**information**’, where the information at any time is the state of the environment, together with the states of all the agents, that are used to define the system’s behavior rules.

For example, consider a colony of ants (a complex system) located on an imaginary grid (its environment). On a few cells of the grid is food (a state of the environment). To collect the food, each ant follows its inherited behavior rule (if it finds food, carry it back to the nest, and lay down a scent trail along the return path; if it comes upon a food scent trail, follow it with a probability that increases with the scent strength; if there is no food or scent nearby, move in a random direction). The information the colony processes at each time step is the state of each cell in the environment grid (the amount of food, and the scent strength). The inherited behavior rule is the colony’s program; the computation result is the distribution of ants on the environment, and the amount of food collected.¹⁸

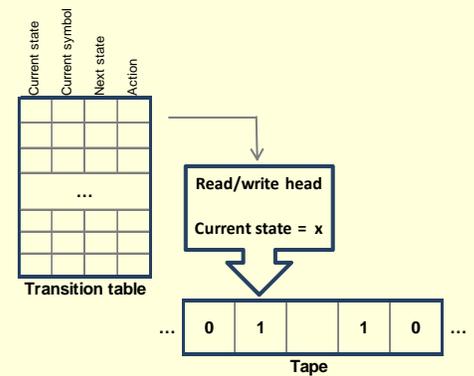
Because it provides a consistent way to analyze the behavior of a complex system, the notion of computation (or information processing) is important. And it can lead to important results: for example, in 1985 the mathematician Alain Lewis proved that perfect rationality, one of the cornerstones of traditional economics, is incomputable. Thus, no matter how sophisticated their behavior rules, agents in a complex system (such as humans) can never be perfectly rational. So, for real-world complex systems such as economic systems, perfect rationality is a fiction.¹⁹

Sometimes, as you will see in Chapter five, a complex system’s computation is equivalent to a Universal Turing Machine (see sidebar).

Universal Turing Machines

In the 1930s, while thinking about Gödel’s incompleteness theorem, Alan Turing wondered if he could construct a machine that not only would perform simple arithmetic, but also would help investigate the limits of what can be computed. In 1937, he described such a machine, now known as a Turing Machine. It consists of three parts:

1. A **tape** that is infinitely long and divided into successive cells, each of which has either a “0”, a “1”, or a blank.
2. A **read/write head** that can move to a particular cell and either read its symbol or write a symbol to it. Associated with the read/write head is a **current state**.
3. A **transition table** with a set of rules that tells the read/write head its next state and its next action (write a symbol, or move one cell left or right), based on its current state and the symbol at the head’s current position.



With an appropriate transition table, a Turing machine can perform any arithmetic or logical function, even one that is infinitely long. Modern computers are finite implementations of a Turing Machine.

Remarkably, Turing proved that there exists a Turing Machine that can reproduce the behavior of any other Turing Machine, including itself. This is a Universal Turing Machine.¹⁷

¹⁷ For a complete description of the Turing Machine, see the Stanford Encyclopedia of Philosophy, at “plato.stanford.edu”.

¹⁸ For an excellent discussion of computation and information processing in complex systems, see Mitchell (2009), chapters 10, 11, and 12.

¹⁹ Lewis (1985)

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

G. BASIC CONCEPTS CONTINUED

2. Non-linear

Physicist Heinz Pagels, one of the early prophets of Complexity Science, wrote, “Life is nonlinear, and so is just about everything else of interest.”²⁰ What did he mean?

A linear relationship (or function f) satisfies the following principles:

$$f(x + y) = f(x) + f(y) \text{ (the principle of additivity), and}$$
$$f(ax) = af(x) \text{ (the principle of homogeneity)}^{21}$$

Change in an independent variable produces a proportional change in a dependent variable. For example, the function $f(x) = 2x$ is linear.

A ‘**non-linear**’ relationship is one that is not linear; change in an independent variable may produce wildly non-proportional change in a dependent variable (see the sidebars). For example, the function $f(x) = x^2$ is non-linear.

Non-linear relationships are important because the computations of most complex systems are non-linear. As examples, the weather, diseases, populations, and financial markets all follow non-linear dynamics. The concept of non-linearity is also important because it helps clarify the difference between Complexity Science and traditional science. The assumption of linearity is the heart of traditional science’s reductionist approach: split a system into parts and analyze the computation of each part; the computation of the whole then equals the sum of the computations of the two parts. Unfortunately, for nearly all complex systems, this approach doesn’t work.

As you study Complexity Science, beware: some authors (especially social scientists) use the term ‘non-linear’ in a different way. By ‘non-linear’, they mean that a system’s behavior rules and relationships give rise to mutually-reinforcing feedback among agents. Even though feedback loops usually produce non-linear system dynamics, the description of ‘non-linear’ above is more common and preferred.

Non-elephants

It is peculiar to refer to most of the world’s functions as ‘non-linear’. As the great mathematician Stanislaw Ulam quipped, “This is like referring to the class of animals that are not elephants as non-elephants.”

Financial markets and yo-yos

“In linear equations, two plus two equals four. Linear equations describe straight lines, discrete phenomena, and an exceedingly small portion of our everyday experience. In nonlinear equations, the effect is not proportional to the cause. The straw that breaks the camel’s back is nonlinear. A small shove can result in a big push. A system evolving in one direction can suddenly veer off in another. Thermometers and bathroom scales are linear. Financial markets yo-yoing between bubbles and crashes are nonlinear. For centuries, scientists reduced the world to linear equations because these were what they could solve. This changed in the late 1970s, with the invention of the personal computer. Its number crunching skills revolutionized physics by allowing scientists to calculate nonlinear equations. It also allowed them to iterate these calculations indefinitely, which sometimes produces the surprising result that two plus two does not equal four.”

Thomas Bass²²

²⁰ Pagels (1988), page 73.

²¹ For all a except complex numbers, additivity implies homogeneity. Note also that there are *two* definitions of ‘linear function’, the one given above, and one from analytic geometry where a linear function is a first-degree polynomial of one variable, the familiar $f(x) = mx + b$. When $b = 0$ and a is not complex, the definitions are equivalent.

²² Bass (1999), pages 65-66.

G. BASIC CONCEPTS CONTINUED

3. Emergence

In Complexity Science literature, the term ‘**emergence**’ is ubiquitous, but its meaning is often muddled.

In their book *Growing artificial societies*, Joshua Epstein and Robert Axtell clearly define emergence as a stable macroscopic or aggregate pattern induced by the local interaction of agents.²⁴ For example, ants and termites following simple behavior rules produce complex patterns of aggregate social behavior; or, as you will see in Chapter six, agents in the artificial society called Sugarscape, following simple trading rules, produce a social system with an aggregate pattern of skewed wealth distribution. In this report, it is with this sense that I use the term ‘emergence’.

However, in Complexity Science literature generally, this term has become muddled almost to the point of uselessness, with many authors using it to describe a result that is surprising or mysterious. One might ask, “Surprising to whom?” Given the term’s arbitrary current usage – and its controversial history – Epstein recently wrote, “I have researched this term more deeply and find myself questioning its adoption altogether.”²⁵

Nevertheless, emergence properly conceived is one of the cornerstones of Complexity Science: aggregate patterns of a complex system arise out of the endogenous interactions of its agents with each other and an environment, without any central controller or other outside influence.

Examples of aggregate patterns are:

- ‘**oscillation**’: Swings in aggregate properties of complex systems are a common emergent pattern. Numbers of predators and prey oscillate; corporations go through business cycles; and financial market prices oscillate.
- ‘**punctuated equilibrium**’: Many complex systems go through long periods of relative stasis, interspersed with brief periods of explosive activity.
- ‘**power laws**’: The pattern of skewed distribution called power law (or, equivalently, ‘Zipf law’ or ‘scale-free distribution’) is covered in Chapter four.

²³ Aaron (1999), page 6.

²⁴ Epstein & Axtell (1996), page 33.

²⁵ Epstein (2006), page 31. This book is one of this report’s *Top ten Complexity Science books*. For an engaging discussion of the term ‘emergence’ and its colorful history, see pages 31-38.

Emergence

In his paper *Predicting the unpredictable*, Eric Bonabeau writes:

“... you first need to understand the concept of ‘emergent phenomena,’ and the best way to do that is by thinking of a traffic jam. Although they are everyday occurrences, traffic jams are actually very complicated and mysterious. On an individual level, each driver is trying to get somewhere and is following (or breaking) certain rules, some legal (the speed limit) and others societal or personal (slow down to let another driver change into your lane). But a traffic jam is a separate and distinct entity that *emerges* from those individual behaviors. Gridlock on a highway, for example, can travel backward for no apparent reason, even as the cars are moving forward.

Emergent phenomena are not just academic curiosities; they lie beneath the surface of many mysteries in the business world. How prices are set in a free market is but one illustration. Why, for example, do employee bonuses and other incentives sometimes lead to reduced productivity? Why do some products – like collapsible scooters – generate tremendous buzz, seemingly out of nowhere, while others languish, despite their multimillion-dollar marketing campaigns? ...

Because of their very nature, emergent phenomena have been devilishly difficult to analyze, let alone predict. Traditional approaches like spreadsheet and regression analyses or even system dynamics ... are currently impotent in analyzing and predicting them. Such approaches work from the top down, ... whereas the behavior of emergent phenomena is formed from the bottom up, starting with the local interactions of the different independent agents.”²³

G. BASIC CONCEPTS CONTINUED

4. Evolution

Evolution is a mechanism of emergence. In Complexity Science, the concept of ‘**evolution**’ is more than Darwin’s theory of biological evolution; it is the way complex systems of all kinds – including businesses, markets, and economic sectors – create behavior patterns that solve hard problems, most commonly problems of survival.

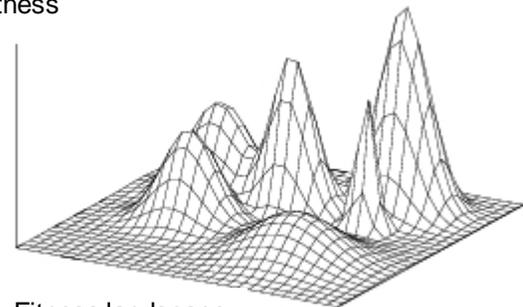
To understand the concept, consider an insurance company competing for survival and market dominance. The company is a complex system, made up of many agents (generally people) in a particular network of relationships, following prescribed behavior rules (many of which are encoded in documents like the company’s mission statement and its business policies and procedures). The company’s purpose is to select from all possible combinations of relationship networks and behavior rules (its ‘**design space**’) the particular combination that will best enable it to survive and dominate. How does it do this?

The design space can be thought of as a 3-dimensional landscape (a ‘**fitness landscape**’), where each point corresponds to one combination of behavior rules and agent relationships. Some combinations will lead to certain failure (low fitness), and some will be winners. But the number of combinations is vast. How does the company search through all the combinations to find an optimal fitness peak?

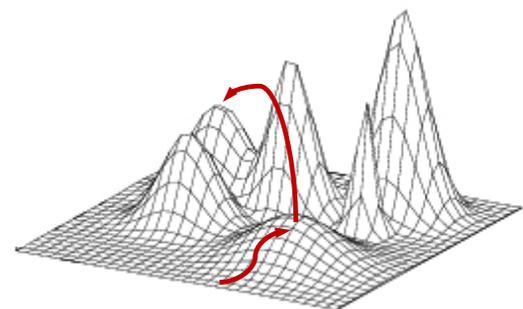
The answer is evolution. Built into the successful company’s behavior rules is an evolutionary algorithm consisting of experimentation and random mutation. Local experimentation enables its objects to try out different combinations of relationships and behavior rules that are similar to current combinations. Some of these will increase fitness, and will be replicated throughout the organization, securing for it a higher position on the fitness landscape. This is the slow walk up the low hill shown in the diagram at right. But the evolutionary algorithm also includes random mutation, wild jumps that can land on much higher fitness levels. Such mutations will also be replicated and increase the company’s ability to thrive.²⁶

Thus evolution is a mechanism of emergence.

Fitness



Fitness landscape



²⁶ For a fascinating account of evolution in Complexity Science, read Beinhocker (2006), chapter nine. This book is one of this report’s *Ten best Complexity Science books*.

G. BASIC CONCEPTS CONTINUED

5. Self-organization

Theoretical biologist and legendary Complexity Science pioneer Stuart Kauffman writes, “The living world is graced with a bounty of order. Each bacterium orchestrates the synthesis and distribution of thousands of proteins and other molecules. Each cell in your body coordinates the activities of about 100,000 genes and the enzymes and other proteins they produce. Each fertilized egg unfolds through a sequence of steps into a well-formed whole called, appropriately enough, an organism. If the sole source of this order is what Jacques Monod called ‘chance caught on the wing,’ the fruit of one fortuitous accident after another and selection sifting, then we are indeed improbable. Our lapse from *paradise ... leaves us spinning around an average star at the edge of a humdrum galaxy, lucky beyond reckoning to have emerged as living forms.”²⁸

But, he is convinced, evolution (‘selection sifting’) is not the only mechanism of emergence. Another is ‘**self-organization**’, the propensity of dynamic systems to organize themselves into complex systems, on their own – without experimentation, mutation, or selection – and seemingly counter to the Second Law of thermodynamics. Stuart Kauffman calls such self-organization ‘order for free’ (see sidebar).

A related concept is ‘**self-organized criticality**’ which we will cover in Chapter five when we study Per Bak’s Sandpile model.

6. Robust

Complex systems are often described as robust or fragile, or both. A ‘**robust**’ complex system is one that manages to survive even when its agents are removed or damaged. For example, the structure and culture of the typical business organization, from the corner store to a multi-national conglomerate, persists, even though it may experience significant personnel turnover.

Similarly, a ‘**fragile**’ complex system is one that can fail if only a few of its agents are damaged. As you will discover in Chapter four, many modern complex systems of interest to actuaries, such as the world financial system, are simultaneously robust and fragile.

Stuart Kauffman

Stuart Kauffman, a theoretical biologist and physician, is a creative and influential complexity scientist with many academic accolades, including a MacArthur ‘genius’ award and a faculty position at the Santa Fe Institute.

In 1965, while in medical school, he thought about light bulbs. What if a hundred light bulbs in a 10x10 grid were wired together such that the state of each bulb (on or off) at time $t+1$ depends on the state of two other randomly selected bulbs at time t . Suppose further that the behavior rule defining the bulbs state at time $t+1$ is selected randomly from the six possible rules (if the other bulbs are both on at time t , the bulb will be on at time $t+1$, etc.). What would happen? Would the hundred lights randomly turn on and off? Would they all freeze into one state?

He programmed his light bulb problem into an IBM computer (with punch cards) and was astounded by the result: most of the bulbs settled into a static on or off state, while about 10 of them fell into a cyclical pattern of recurrent blinking. The result was a major experience in his life. About this result, many years later he said, “I’m still deeply proud of that. I’m still stunned that if you make a random network with light bulbs and everybody has two inputs per light bulb, and otherwise you make everything at random, the thing behaves with order. Still blows me away! Thirty-seven years later, still blows me away!”²⁷

That experience, and many follow-on simulations and experiments led him to conclude that evolution is not the only mechanism leading to emergence in complex systems, and is not even the most important mechanism. Rather, agents self-organize. We get “order for free”.

²⁷ For a more detailed description of this experience, and of Stuart Kauffman, see Regis (2003), chapter two.

²⁸ Kauffman (1995), page 71.

G. BASIC CONCEPTS CONTINUED

7. Edge of chaos

The term ‘**edge of chaos**’ may have more usefulness as a metaphor than as an essential Complexity Science concept. This term signifies that the most interesting complex systems are those closer to random systems (ie, nearer to the edge of chaos) than to simple systems. For a more colorful description of this concept, see the sidebar.

In their book *Complex adaptive systems*, John Miller and Scott Page propose that the ‘edge of chaos’ concept has both a weak and a strong form, and that the strong form may be incorrect.

- weak form: the most interesting and productive complex adaptive systems lie somewhere in the space between simple systems and random systems.
- strong form: (the form most authors cite) the most productive complex systems lie very close to (at the edge of) random systems.

Miller and Page write, “One hypothesis is that adaptive systems will have a bias toward emphasizing simple structures that resist chaos over more complicated ones that handle difficult situations. There are two reasons for this hypothesis. The first is that simple structures are likely to be easier to find and maintain. ... The second justification for the hypothesis is that systems that are fragile are very risky in terms of rewards, and adaptive systems tend to be risk averse. While being able to handle delicate situations appropriately on occasion might result in large rewards, there is also a chance that it will lead to large losses. ... The strong-form hypothesis – namely, that adaptive systems congregate at a narrow edge where slight changes in their behavior lead to chaos or frigidity – is harder to justify.”²⁹

Stuart Kauffman concurs: “It is far too early to assess the working hypothesis that complex adaptive systems evolve to the edge of chaos. Should it prove true, it will be beautiful. But it will be equally wonderful if it proves true that complex adaptive systems evolve to a position somewhere in the ordered regime near the edge of chaos. Perhaps such a location on the axis, ordered and stable, but still flexible, will emerge as a kind of universal feature of complex adaptive systems in biology and beyond.”³¹

Spontaneous, adaptive, alive

In *Complexity*, Mitchell Waldrop writes:

“This balance point – often called the *edge of chaos* – is where the components of a system never quite lock into place, and yet never quite dissolve into turbulence, either. The edge of chaos is where life has enough stability to sustain itself and enough creativity to deserve the name of life. The edge of chaos is where new ideas and innovative genotypes are forever nibbling away at the edges of the status quo, and where even the most entrenched old guard will eventually be overthrown. The edge of chaos is where centuries of slavery and segregation suddenly give way to the civil rights movement of the 1950s and 1960s; where seventy years of Soviet communism suddenly give way to political turmoil and ferment; where eons of evolutionary stability suddenly give way to wholesale species transformation. The edge of chaos is the constantly shifting battle zone between stagnation and anarchy, the one place where a complex system can be spontaneous, adaptive, and alive.”²⁹

²⁹ Waldrop (1992), page 12.

³⁰ Miller & Page (2007), page 140.

³¹ Kauffman (1995), page 91.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

H. TOOLS

To model the behavior of complex systems, complexity scientists employ many tools. Although most have been assimilated from other fields, such as physics and mathematics, Complexity Science applies them in new ways and to new complex systems. A common characteristic of all the tools is that they are computer-intensive.

Some tools are used to set up the model structure, including the structure of agent relationships and agent behavior rules; and others are for real-world data analysis to develop agent attributes or behavior rules, or to validate model results.

Model structure

The primary tool to set up a model's structure is the '**agent-based model**'. Chapter two describes this tool, and throughout the report are many examples of its use.

To organize agent relationships, complexity scientists use '**graphs**' and '**networks**'. Chapter four describes the use of these tools.

Agent behavior rules can be based on a simple 'if-then' structure, or more complicated. Common tools that complexity scientists use to organize behavior rules are '**game theory**', '**genetic algorithms**', '**heuristics**', and '**neural networks**'. Chapter two covers these.

Data analysis

To analyze real-world data, complexity scientists employ many tools that are familiar to actuaries, such as time series analysis, data visualization, and data mining. They also employ tools that may be unfamiliar, such as:

- '**controlled experiment**': Controlled experiments, especially behavioral economics experiments, that uncover real-world behavior rules that agents follow. These are covered in Chapter five.
- '**multi-dimensional histograms**': Histograms showing frequencies of more than one system state over time.
- '**pattern matching algorithms**': Algorithms that find data patterns with potential near-term predictive value.
- '**phase space diagrams**': Multi-dimensional diagrams that show the possible states of a system, with each state corresponding to a point on the diagram.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

I. KEY INSIGHTS

This section summarizes many of Complexity Science's key insights. These help us better understand the nature of complex systems, and have led to many practical results.

- **Emergence from simple rules.** Complex systems with intricate and interesting behavior can emerge from agents following simple behavior rules. This is one of the key insights from Stephen Wolfram's work with cellular automata, which we will cover in Chapter five.
- **Impossibility of long-term prediction.** Predicting the long-term behavior of complex systems is impossible. This is the crux of Wolfram's Principle of Computational Equivalence that we'll study in Chapter five. Indeed, it is a theme running throughout Complexity Science, and one of the five major themes of this report.
- **Self-organization.** Agents can start out in complete disarray, and by merely following simple behavior rules generate a complex system that is highly organized. They do this without any central control mechanism. This is one of the central insights of Stuart Kauffman and Per Bak that we will explore in Chapter five. Some authors call this result 'order for free' or 'spontaneous order'.
- **Ubiquitous power laws.** Complex system properties often follow power laws. This is one of the common patterns of emergence. We will encounter this result in Chapters four, five, and six.
- **Punctuated equilibrium.** Another common pattern of emergence is punctuated equilibrium, long periods of relative stasis, interspersed with brief periods of explosive productivity.
- **Robust and fragile.** Complex systems are often simultaneously robust and fragile. For example, as you will see in Chapter four, many real-world networks such as the Internet are robust to random attack, but fragile to focused attack.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

J. PRACTICAL RESULTS

In each of Chapters four through seven, there is a section titled *Practical applications*. Each section gives examples how Complexity Science has produced results of significant practical importance. These results are listed below.

Chapter four (Networks) examples:

- Explain the 2000 worldwide airline crisis.
- Show how to make the world airline network less vulnerable to terrorist attack.
- Show how to manage the spread of AIDS.
- Help make the world financial network more resilient.
- Map the resilience of organizations (for enterprise risk management).

Chapter five (Cellular automata) examples:

- Explain financial market price fluctuations.
- Show the aggregate behavior of infectious diseases, including power-law distribution and cyclical dynamics.
- Show the impact of media and population density on public opinion.
- Help to understand the impact of patient choice on mortality and complications of complex surgical procedures.
- Explain why the 1961 change in the US minimum Social Security retirement age had an unexpected impact.
- Show the potential impact of a new guaranteed income variable annuity product, to help determine its pricing.

Chapter six (Artificial societies) examples:

- Explain why a society's wealth ends up in a skewed distribution.
- Explain why in real economies price and quantity traded do not correspond to supply and demand curves.
- Help develop national strategy to counter bioterrorism.
- Determine the best screening strategy for type 2 diabetes.

Chapter seven (Serious games) examples:

- Help develop effective business strategies for insurance and re-insurance companies.
- Explain property and casualty re-insurance price cycles.
- Train students how to manage an insurance company.
- Enable consumers to design health insurance plans.

J. PRACTICAL RESULTS CONTINUED

Following are additional practical results of Complexity Science:

- **Remaking economics.** One of the most far-reaching practical results of Complexity Science is that it is remaking the field of economics. Eric Beinhocker writes about this in his book *The origin of wealth*.³³
- **Understanding the impact of regulations and policy change.** Complexity Science is helping regulators and policy makers understand the impact of proposed changes. For example, to avoid repeating the 2000 disaster when Enron and other companies manipulated energy supplies and prices, several US states now use agent-based models to test complex electricity market designs before implementation.³⁴
- **Generating business savings.** Applying Complexity Science methods helped Citibank uncover over \$200 million of previously unidentified exposure for delinquent credit card payments, Proctor and Gamble save 22 percent in distribution expense, DuPont save \$500 million annually in manufacturing expense, the Internal Revenue Service improve its fraud detection capability by 8,000 percent, Nasdaq understand the impact of changing stock prices from fractions to decimals, Hewlett-Packard anticipate how changes in its hiring strategy would affect its corporate culture, Société Générale determine operational risks of its asset management group, and Southwest Airlines save \$10 million yearly in labor costs.³⁵

Southwest Airlines approached Stuart Kauffman to ask his help to improve the company's freight distribution process. Using agent-based simulation models, Kauffman and his team of complexity scientists discovered that the airline's cargo handlers (the complex system's agents) followed a behavior rule (that Kauffman dubbed 'the hot potato rule') leading to an emergent pattern of backlogged freight. Kauffman found, and Southwest implemented, a better behavior rule, which resulted in considerable savings.³⁶

- **Automating financial trading.** Complexity Science tools and insights applied to financial markets have resulted in vast profits (see the sidebar).

The Prediction Company

Two childhood friends from Silver City, New Mexico – Doyne Farmer and Norman Packard – both became complexity scientists, and together founded one of the most successful Complexity Science businesses, The Prediction Company.

Located in Santa Fe, The Prediction Company uses Complexity Science tools and insights to build financial market trading systems. When they started the company in 1991, the two scientists knew virtually nothing about trading financial instruments. By 2001, the Prediction Company's 'black box' trading systems had become the finest in the world. In 2005, the company was purchased by one of the largest banks, Union Bank of Switzerland.

The company's systems automate the financial trading process. They crunch through terabytes of worldwide financial data, using armies of sophisticated pattern recognition and learning programs – like genetic algorithms and neural networks – to find data patterns that are likely to reappear. They then automatically place buy and sell orders to take advantage of the patterns. Even though their systems only produce small near-term predictive advantages over other systems, because their trade volume is high, they produce vast profits.³²

³² To learn more about The Prediction Company, read Bass (1999).

³³ Beinhocker (2006).

³⁴ Economist Leigh Tefatson of Iowa State University (in Ames, Iowa) has led the development of the agent-based model known as the *Ames wholesale power market test bed*.

³⁵ Smith & Segre-Tossani (2003) and Bonabeau (2002).

³⁶ Read Regis (2003) for more information about this and other practical results of Complexity Science.

K. ISSUES AND FUTURE DIRECTION

In 1996, John Horgan wrote dismissively about complexity scientists, “They will make incremental advances, such as extending the range of weather forecasts or improving the ability of engineers to simulate the performance of jets or other complex technologies. But they will not achieve any great insights into nature – certainly none comparable to Darwin’s theory of evolution or quantum mechanics. They will not force any significant revisions in our map of reality or our narrative of creation.”³⁸

Horgan’s main complaints about Complexity Science were:

- that it employed terminology amounting to little more than metaphor, and
- that its computer models bore little resemblance to reality.

As you have seen, his complaint about terminology is apt still. But, his complaint about computer models has turned out to be unfounded: through their computer models, complexity scientists such as Stephen Wolfram, Joshua Epstein, and Stuart Kauffman are significantly revising our map of reality.

Although Complexity Science has already achieved more than Horgan thought it ever would, it is still plagued by the lack of a conceptual foundation. Its vocabulary is still imprecise, it has yet to develop a logical foundation, and there are basic questions that remain unanswered (see sidebar).

Nevertheless, complexity scientists are providing rigorous theory and tools to illuminate our understanding of natural and social complex systems, and to help us better manage our complex financial, insurance, pension, health, and political systems. I hope that complex systems actuaries will add to this store of knowledge and tools.

(In the last few years, as Complexity Science has matured, Horgan and other detractors have been silent.)

Open questions

In their book *Complex adaptive systems*, John Miller and Scott Page list several open questions for Complexity Science.³⁷ Following are extracts:

What does it take for a system to exhibit complex behavior? What is it about interacting agents that leads to complex behavior? Is there some behavioral threshold that must be breached before complexity can arise?

Is there an objective basis for recognizing emergence and complexity? If you put a frog in a blender and turn it on, there is only a macabre interest in the resulting chemical soup. If, however, you start with a chemical soup and run the blender backwards, and out of the froth pops a fully formed frog, then something rather different has happened. Is there some easy and reliable way to separate out these two experiences? Of course, this would matter little if we weren’t seeing so many frogs popping out of the froth of both nature and our models. ... How do we separate complex systems from merely complicated ones?

What mechanisms exist for tuning the performance of complex systems? Insofar as real social systems behave according to the laws of complex adaptive social systems, what policies can be used to direct the outcomes of these systems?

What makes a system robust? What does it take for a system to persist in the face of external changes? Alternatively, we could frame the question as uncovering the factors that make a given system brittle.

³⁷ Miller & Page (2007), Appendix A.

³⁸ Horgan (1996), pages 225-226.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

L. EXERCISES

1. Check the appropriate cells to describe each item:

Item	Set	System	Dynamic system	Simple system	Random system	Complex system	Complex adaptive system
a ball, a blanket, and a bird							
a traffic light							
air molecules in a room							
a cloud							
the weather							
the solar system							
a soccer team							
a termite colony							
the U.S. healthcare system							

For each system, identify its objects. If there are relationships among the objects, behavior rules, or an environment, also identify these.

2. Think of members of an orchestra on a stage as agents in a dynamic system, each of which can play one sound (its state) at any time. Then consider a music score as the agents' behavior rule.

- What kind of score would turn the musicians into a simple system, a random system, or a complex system?
- Can you think of a score that would turn the agents into a complex adaptive system?

3. Using the example of an orchestra in the previous example, show how it performs a computation. Identify the information, the program, the computation, and the computation result. Does this computation produce an emergent property? Is the orchestra a universal computer?

4. Contrast how a complex systems actuary and a traditional actuary might approach the following problem: It is well-known that physicians change their diagnostic and prescribing practices to adapt to changes in insurance reimbursement policies. Your employer, a health insurance company, has asked you to determine the expected impact on company profits from implementing a new policy that reduces to zero its reimbursement for a popular diagnostic test.

Complexity science – an introduction (and invitation) for actuaries

Chapter one: Complexity Science continued

M. TO LEARN MORE

To learn more about the new Complexity Science, you may enjoy watching the series of NOVA videos titled “Emergence – complexity from simplicity, order from chaos”³⁹. Although now dated, Mitchell Waldrop’s book *Complexity* (one of this report’s *Top ten Complexity Science books*) is an excellent introduction to the field.⁴⁰

To learn more about the difficulties of prediction, you may enjoy the book *The future of everything – the science of prediction* by David Orrell.⁴¹

N. REVIEW AND A LOOK AHEAD

This chapter introduced Complexity Science, including its development, hallmarks, concepts, tools, key results, and issues.

The next chapter describes the model structure used to simulate complex systems: the agent-based model.

³⁹ NOVA (2007)

⁴⁰ Waldrop (1992)

⁴¹ Orrell (2007)

CHAPTER TWO: AGENT-BASED MODELS

The agent-based computational model – or artificial society – is a new scientific instrument. It can powerfully advance a distinctive approach to social science, one for which the term ‘generative’ seems appropriate.

Joshua Epstein²

A. INTRODUCTION

The heart of Complexity Science is agent-based modeling. This chapter introduces you to ‘**agent-based models**’ (also called ‘**computational models**’ and ‘**multi-agent models**’³), describes their purpose and characteristics, compares them to actuarial modeling, and – most usefully – discusses how you can build your own.

B. PURPOSE

Many actuaries believe the primary purpose of modeling is prediction, to predict policyholder events, economic measures like interest rates, and healthcare expenditure trend rates. But prediction is only one of the many purposes of agent-based models, and – especially considering the futility of trying to predict the long-term behavior of complex systems – not the most important.

Joshua Epstein provides sixteen good reasons other than prediction to build agent-based models (see sidebar). Let’s look at a few that are particularly relevant for actuaries:

Explanation

For complex systems, explanation is more important than prediction. For example, because agent-based models have helped us better understand relationships between epidemic dynamics and underlying population configurations, we can now implement better containment strategies, even though we are powerless to predict an epidemic’s course.

Similarly, explanatory models that increase our understanding of actuarial risks can help us better manage them, even if we cannot predict their incidence or consequences.

Why model?

“The modeling enterprise extends as far back as Archimedes; and so does its misunderstanding. ... The first question that arises frequently – sometimes innocently and sometimes not – is simply, ‘Why model?’. ... my favorite retort is, ‘You *are* a modeler.’ Anyone who ventures a projection, or imagines how a social dynamic – an epidemic, war, or migration – would unfold is running *some* model. But typically it is an *implicit* model in which the assumptions are hidden, their internal consistency is untested, their logical consequences are unknown, and their relation to data is unknown. The choice, then, is not whether to build models; it’s whether to build *explicit* ones. ...

No sooner are these points granted than the next question inevitably arises: ‘But can you predict?’ For some reason, the moment you posit a model, prediction ... is reflexively presumed to be your goal. Of course, prediction might be a goal ... But, more to the point, I can quickly think of *16 reasons other than prediction ... to build a model ...*

1. Explain (very distinct from predict)
2. Guide data collection
3. Illuminate core dynamics
4. Suggest dynamical analogies
5. Discover new questions
6. Promote a scientific habit of mind
7. Bracket outcomes to plausible ranges
8. Illuminate core uncertainties
9. Offer crisis options in near-real time
10. Demonstrate tradeoffs/ suggest efficiencies
11. Challenge the robustness of prevailing theory through perturbations
12. Expose prevailing wisdom as incompatible with available data
13. Train practitioners
14. Discipline the policy dialogue
15. Educate the general public
16. Reveal the apparently simple (complex) to be complex (simple)”

Joshua Epstein¹

¹ J. M. Epstein (2008)

² Epstein (2006), page 5.

³ In current usage, the term ‘multi-agent model’ has various meanings that can be confusing. Sometimes it is used as a synonym for ‘agent-based model’. But the term can also mean a model in the field of multi-agent systems, a field mainly concerned with robot interactions. The term can also be used to mean a subset of agent-based models in which agents are heterogeneous.

Complexity science – an introduction (and invitation) for actuaries

Chapter two: Agent-based models continued

B. PURPOSE CONTINUED

Data collection

In order to explain a phenomenon of interest, a model can help clarify the data we should collect. Such a model would contribute to an iterative process, for data can also guide model design.

Confidence intervals

Even in cases where we cannot predict, models can help us understand plausible outcome ranges. Using sensitivity analysis, we can explore a range of parameters to identify salient uncertainties, regions of robustness, important thresholds, and possible outcomes.

Real-time crisis management

In complex systems, crises will strike at unexpected times, in unexpected ways. This is one of the lessons of Complexity Science. Although agent-based models won't improve our ability to predict the behavior of complex systems, they can help actuaries alert their employers and clients about the kinds of surprises that may emerge, and to suggest small interventions that will produce significant returns. This ability is particularly important for actuaries employed in enterprise risk management.

For example, in Chapter four you will learn how network analysis:

- can help mitigate the impact of an infectious disease, even though we cannot predict its course, and
- can make networks such as the airline system and the Internet more resilient to attack, even though we cannot predict when an attack will occur.

Training

Models can help actuaries and executives better understand and develop intuitions about the behavior of the complex systems in which they operate. Serious games are particularly suited for this (see Chapter seven).

Policy support

Policy proposals are often developed by people who do not understand the complex systems for which they are proposing change. By modeling the behavior of such systems, actuaries can support policy development.

Public education

The public does not understand what actuaries do. Simple models can help people understand actuarial work.

C. KEY CHARACTERISTIC

The key characteristic of an agent-based model is that it is bottom-up. The model specifies the agents that comprise a system (see sidebar) and the relationships among agents; it may also specify agent behavior, agent interactions with an environment, and even the involvement of a human player. But that's all; the model does not impose top-down criteria.

Based on the model's specification, the computer then runs myriad agent interactions, creating the system's behavior from the bottom up. Overall system behavior patterns gradually emerge, as in reality.

D. STRUCTURE

No matter whether agent-based models are implemented as Excel spreadsheets, advanced Java models, or sketches on a napkin, their structure generally includes the following elements:

Agents and their attributes

Databases or iterated functions define the number of agents and their individual attributes.

Agent relationships

Functions or databases define how agents are related. For example, all of the Excel models in Chapter five (cellular automata) use one- or two-dimensional grids on a spreadsheet to define agent relationships.

Agent behavior

One or more functions define the rules for how the agents behave with one another and with the environment. The next section of this chapters treats behavior rules in detail.

Environment

If agents can move or interact with an environment, the model includes an environment. It can be as simple as a one-dimensional grid, or as complex as a realistic topographic map.

Time

A model incorporates a method to organize sequential agent behavior, typically as time steps or 'ticks'.

Analysis

A model usually includes analyses of its results. For example, Excel models in Chapter five include graphs of model results.

User interface

Models typically include a way for users to enter and change parameters, as well as a visual display of agents as they interact.

Agents

An '**agent**' is the part of a model that represents an actor within a system. It might represent a person, an organization, an economic sector such as health care, a governmental entity, or a country. Agent-based models typically include many such agents, often thousands or millions.

Following are agent characteristics.

Hierarchical

An agent can include other agents, just as organizations include people.

Local

Agents act locally. They generally only interact with other agents within a relatively small neighborhood.

Autonomous

Agents take independent action to attain their goals.

Boundedly rational

Agent behavior rules are based on real-world behavior, which is far from perfectly rational; people and organizations generally act based on limited knowledge and simple – often illogical – heuristics; they are what behavioral economists call 'boundedly rational'.

Adaptive

Agents can learn to adapt to their environment, and change their behavior based on what they learn.

Heterogeneous

Agents can be quite different from one another and follow different behavior rules. In fact, for a system to be robust, agents *must* be diverse. As you will see in Chapter four, breakdowns in agent diversity can lead to correlated behavior and massive disruptions, such as the recent worldwide financial meltdown.

E. AGENT BEHAVIOR RULES

Agent-based models are the heart of Complexity Science, and at the heart of agent-based models are agent behavior rules. An agent's behavior rule determines how it changes its state in response to the states of other agents and the environment.

This section describes several types of behavior rules, and shows how you can develop such rules for your Complexity Science models.

1. Behavior rule types

Simple if-then rule

The simplest, and most common, behavior rule is the simple if-then rule. As an example of a behavior rule for a physician agent: **If** the result of a total cholesterol test for a patient agent is greater than 240 mg/dl, **then** order a fasting lipid profile test. A simple if-then rule is often expressed as a '**transition table**', examples of which you will find below and in Chapter five.

Genetic algorithm

Inspired by the mechanics of biological evolution, genetic algorithms were created by John Holland (see sidebar). They enable agents to efficiently search complex fitness landscapes for optimal solutions. To employ a genetic algorithm, an agent first uses a fitness function to compare the relative values of competing solutions (points on the fitness landscape). It then selects the best solutions and combines them into a new set of solutions (new points on the fitness landscape) which it tests again. Repeating such a process, it iteratively evolves an optimal solution.

In her book *Complexity*, Melanie Mitchell gives an entertaining example showing how genetic algorithms work⁵: Suppose an agent is a near-sighted robot janitor – named Robby – assigned to pick up empty soda cans on a 10x10 grid. He uses a genetic algorithm to find an optimal (ie, shortest) path to pick up the cans. Because Robby is near-sighted, he can only see the contents of immediately adjacent grid cells to its north, east, south, and west. And he can only take seven actions: move one cell to the north, east, south, or west, move one cell in one of the four directions randomly (based on a randomly-generated number), pick up a can, or do nothing.

John Holland

John Holland, a father of Complexity Science and creator of the genetic algorithm, is the world's first PhD in computer science (from the University of Michigan).

Mitchell Waldrop describes one of the first lectures that Holland gave at the Santa Fe Institute, in 1987:

"He proved to be a compact, sixtyish Midwesterner with a broad, ruddy face that seemed fixed in a perpetual grin, and a high-pitched voice that made him sound like an enthusiastic graduate student. ...

First, he said, each of these [complex adaptive] systems is a network of many 'agents' acting in parallel. ... Furthermore, said Holland, the control of a complex adaptive system tends to be highly dispersed. ... If there is to be any coherent behavior in the system, it has to arise from competition and cooperation among the agents themselves. ... Second, said Holland, a complex adaptive system has many levels of organization, with agents at any one level serving as the building blocks for agents at a higher level. ... Furthermore, said Holland – and this was something he considered very important – complex adaptive systems are constantly revising and rearranging their building blocks as they gain experience. ... Third, he said, all complex adaptive systems anticipate the future. ... every complex adaptive system is constantly making predictions based on its various internal models of the world ... it's no wonder that complex adaptive systems were so hard to analyze with standard mathematics. ...

Holland's ideas produced a shock of recognition, the kind that made more ideas start exploding in your own brain".⁴

⁴ Waldrop (1992), pages 144-148. You may enjoy Holland's lecture about modeling complex adaptive systems: Holland (2008).

⁵ Mitchell (2009), pages 129-142.

E. AGENT BEHAVIOR RULES CONTINUED

1. Behavior rule types *continued*

*Genetic algorithm *continued**

Robby always starts at cell (0,0), and during each cleaning session he can take only 200 actions. Each action meets with reward or punishment: if he is in the same cell as a can and picks it up, he is rewarded 10 points; but if he bends down to pick up a can on a cell where there is no can, he is fined 1 point. If he bumps into a wall, he is fined 5 points. To maximize his reward, Robby wants to pick up as many cans as possible without bending down unnecessarily or crashing into walls.

Each solution on Robby’s fitness landscape is a strategy that Robby can follow to carry out his 200 actions. The best solution is the one that maximizes Robby’s points.

The set of all situations that Robby could encounter can be listed as in the first five columns of the following table:

Current state of environment in cells relative to Robby					action
north	south	east	west	current	
empty	empty	empty	empty	empty	move north
empty	empty	empty	empty	can	pick up can
...
empty	empty	empty	empty	can	move north
wall	empty	empty	empty	empty	move north

There are about 250 possible situations. If for each situation, there is in the last column of the table a corresponding action that Robby can take, then the table is one of Robby’s possible strategies, or one solution in his fitness landscape. Because Robby can take any one of seven actions for each situation, the potential number of solutions (the fitness landscape) is quite large – about 7^{250} . Among all these possibilities, Robby wants to find a strategy that maximizes his points, for any random arrangement of cans on the grid.

Assuming that there are exactly 250 situations, and that the situations are always listed in the same order in the table, each strategy can be represented as a series of 250 numbers from 0 to 6, where each number corresponds to one of Robby’s seven possible actions. For example, one strategy might be:

062454344 ... 533210 (a string of 250 numbers)

E. AGENT BEHAVIOR RULES CONTINUED

1. Behavior rule types *continued*

*Genetic algorithm *continued**

For purposes of the genetic algorithm, we can think of each number in the string as a gene, and the entire string as the DNA of an ‘individual’.

To evolve an optimum strategy, or individual, the genetic algorithm does the following:

1. It generates a random set of 100 strategies (a set of 100 different individuals).
2. For a random placement of cans, Robby then follows each strategy 200 times (corresponding to his maximum of 200 actions) and along the way keeps a tally of his points. The number of points is the strategy’s (individual’s) ‘fitness’.
3. He repeats the previous step 1,000 times, and keeps track of the average fitness of each of the 100 strategies (individuals).
4. He repeats the following steps until he generates a new population of 100 strategies (individuals):
 - a. Randomly choose two ‘parent’ individuals, with the probability of selection proportional to individual fitness.
 - b. Mate the two parents to create two children. To do this, randomly choose a position to split the two number strings, and form one child by taking the numbers before that position from parent A and after that position for parent B, and vice versa for the second child.
 - c. Mutate the children’s DNA. With a small probability, choose one or more numbers and replace them with a randomly generated number between 0 and 6.
 - d. Put the two new children in the new population of individuals (a new generation).
5. Return to step 2.

The algorithm repeats this process multiple times to evolve an optimum strategy (individual) with the highest fitness.

To test Robby’s evolved strategy, Mitchell did her best to develop a strategy on her own, using human judgment (Mitchell has a PhD in computer science, and is a respected complexity scientist). She then compared her strategy with Robby’s over 10,000 cleaning sessions. Mitchell’s average score was 346, while Robby’s average was 483. A perfect score would have been 500.

Because of its effectiveness, the genetic algorithm is often used to model adaptive agents in a complex adaptive system.

E. AGENT BEHAVIOR RULES CONTINUED

1. Behavior rule types continued

Artificial neural network

Artificial neural networks (ANN) were also inspired by biological mechanics, specifically the neural system of the human brain. ANN are useful when an agent needs to learn from past patterns of its complex system in order to determine its future actions, just as humans use their brains to learn and adapt.

To understand the motivation for ANN, let's briefly look at the human brain. It consists of nerve cells called neurons, each with four parts:

- **cell body:** the part of the cell that takes input from other neurons and sends output to other neurons.
- **dendrites:** fibers connected to the cell body that receive input from other neurons
- **axon:** a fiber connected to the cell body that delivers the neuron's output to other neurons.
- **synapse:** where a dendrite fiber from one neuron connects to the axon of another neuron.

The human brain has about ten billion (10^{10}) neurons and about 10^{14} synapses (the average neuron is connected to other neurons with about 10^4 synapses). Neurologists have discovered that the brain learns by changing the strength of synaptic connections according to the number of times the synapses are stimulated. ANN are based on a similar principle.

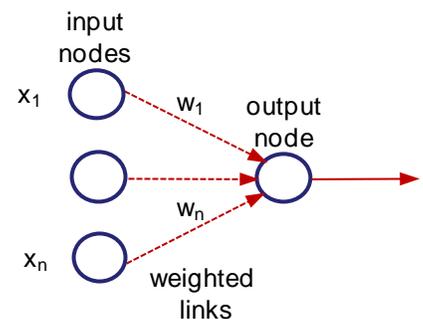
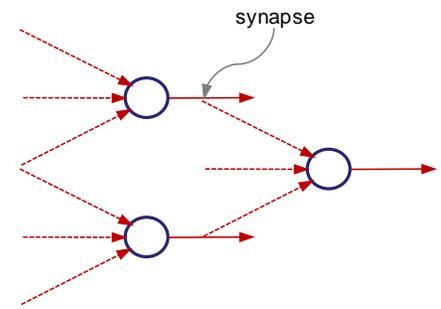
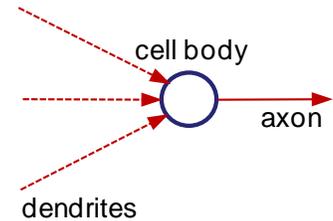
To understand how an ANN works, let's first explore a simple ANN known as a perceptron. The perceptron consists of three parts:

- **output node:** the part of the perceptron that computes the output.
- **input nodes:** the part that collects input.
- **weighted links:** the connections between input nodes and the output node, with each link associated with a weight.

The output that the output node computes is:

$$y = f(x_1 w_1 + \dots + x_n w_n - t)$$

where $(x_1 \dots x_n)$ are inputs, $(w_1 \dots w_n)$ are weights, t is a 'bias factor', and $f()$ is an 'activation function'.



E. AGENT BEHAVIOR RULES CONTINUED

1. Behavior rule types continued

Artificial neural network continued

As in the brain, a perceptron learns by adjusting the weights of its links until the output fits the underlying data.

For example, suppose the agent’s underlying historical data is given by the following table of eight samples:

n	x ₁	x ₂	x ₃	y
1	0	0	0	-1
2	0	0	1	-1
3	0	1	0	-1
4	1	0	0	-1
5	0	1	1	1
6	1	0	1	1
7	1	1	0	1
8	1	1	1	1

Suppose further that its output function f is simply the sign function (producing -1 if its argument is negative, and $+1$ otherwise), and that $t = 0.4$.

To determine an optimal set of weights, the agent does the following:

1. Initializes its weights w_1 , w_2 , and w_3 with random values between 0 and 1.
2. Computes the expected output \hat{y}_n for each historical sample based on the current weights.
3. Updates the weights with the following formula:

$$w_i^{k+1} = w_i^k + \sigma (y_n - \hat{y}_n^k) x_{in}$$

where the superscript k refers to the iteration number, x_{in} is the i th input item for the n th sample, and σ is a parameter called the ‘learning rate’ which for this example we can set equal to 0.5.

4. The agent returns to step 2 and iterates until the weights converge, or until $y_n - \hat{y}_n^k$ reaches the desired degree of precision.

For this example, all the weights converge to 0.3.

E. AGENT BEHAVIOR RULES CONTINUED

1. Behavior rule types continued

Artificial neural network continued

A real ANN is more complicated than a perceptron: it may add several intermediate layers between the input and output. These additional layers are called ‘hidden layers’, and their nodes are ‘hidden nodes’. For a real ANN, the activation function may be any type, including linear, logistic, or hyperbolic tangent functions.

Multilayer ANN are called ‘universal approximators’, because they can approximate any target function.

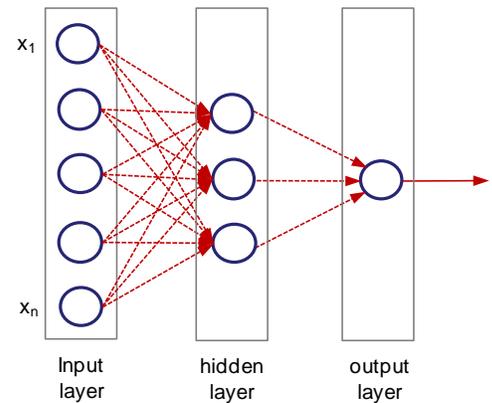
Game theory

In a complex social system, agents usually need to follow behavior rules that incorporate cooperation and competition with one another. To construct such behavior rules, the structure of game theory is useful.

The most common structure in game theory is the ‘prisoner’s dilemma’. It is so common and important that Robert Axelrod, a Complexity Science pioneer, wrote “The two-person iterated Prisoner’s Dilemma is the *E. coli* of the social sciences, allowing a very large variety of studies to be undertaken in a common framework. It has even become a standard paradigm for studying issues in fields as diverse as evolutionary biology and networked computer systems. Its very simplicity has allowed political scientists, economists, sociologists, philosophers, mathematicians, computer scientists, evolutionary biologists, and many others to talk to each other.”⁶

To understand the prisoner’s dilemma, consider two agents who are arrested for a crime (that they did commit).⁷ Because the police have inadequate evidence to convict either, they tell each suspect:

- if he rats against the other, he will receive a reward and will be released, provided the other suspect does not rat;
- if both agents rat against each other, each will go to jail but with a reduced sentence;
- if he keeps quiet and the other agent rats, he will go to jail for a long time;
- if both keep quiet, both will go free.



⁶ Axelrod (1997), page xi.

⁷ There are many variations of the prisoner’s dilemma, with differing rewards.

E. AGENT BEHAVIOR RULES CONTINUED

1. Behavior rule types continued

Game theory continued

The choices in the prisoner’s dilemma can be represented as a ‘payoff matrix’. Each matrix cell contains a pair of payoffs (x,y), where the first of the pair is Agent B’s payoff and the second is Agent A’s. If both keep quiet, both go free and the payoff is (1,1); if both rat, each goes to jail for a reduced term, and their payoff is (-2,-2); if Agent A rats and Agent B keeps quiet, then their payoff is (-5,1) because Agent B goes to jail for a long time and Agent A goes free. Clearly, mutual cooperation is the best outcome (the total payoff is $1 + 1 = 2$). But would you keep quiet?

		Agent A	
		keep quiet	rat
Agent B	keep quiet	(1,1)	(-5,1)
	rat	(1,-5)	(-2,-2)

The dilemma is whether to follow your narrow self-interest (rat and hope the other doesn’t) or cooperate for greater mutual gain (keep quiet and hope the other does also). This dilemma arises in situations as diverse as marriage, business strategy, combat, and nuclear arms control. Although our lives depend on cooperation, the reality is that people usually choose their narrow self-interest over the common good, and end up with a sub-optimal total payoff. Interestingly, for a single-round game such as the classic prisoner’s dilemma, self-interest is actually the most rational choice.

Life’s dilemmas are usually not found in single-round games. For example, marriage, business relationships, and even wars usually have many rounds. What is the best strategy for a many-round game, the so-called ‘iterated prisoner’s dilemma’?

In 1987, Robert Axelrod published the answer. Based on suggestions from John Holland (both were at the University of Michigan), he set up a computer simulation to pit agents against each other for thousands of iterations of the prisoner’s dilemma, and used a genetic algorithm for them to evolve an optimal strategy. The result? The most commonly evolved optimal strategy was ‘tit-for-tat’: cooperate the first time, and thereafter do what the other agent did in the game’s last turn.⁸

Thus, the structure of game theory, combined with the genetic algorithm, evolved an effective behavior rule for agents in many social circumstances. You can also use the structure of game theory to develop agent behavior rules.

⁸ Axelrod (1997)

E. AGENT BEHAVIOR RULES CONTINUED

1. Behavior rule types continued

Heuristics

In his book *Predictably irrational*, Dan Ariely writes, “For a long time, economists have maintained that human behavior and the functioning of our institutions are best described by the rational economic model, which basically holds that man is self-interested, calculating, and able to perfectly weigh the costs and benefits in every decision in order to optimize the outcome. But in the wake of a number of financial crises, from the dot-com implosion of 2000 to the subprime mortgage crisis of 2008 and the financial meltdown that followed, we were rudely awakened to the reality that psychology and irrational behavior play a much larger role in the economy’s functioning than rational economists (and the rest of us) had been willing to admit. ... If the rational economic approach is not sufficient to protect us, what are we supposed to do? What models should we use? Given our human fallibilities, quirks, and irrational tendencies, it seems to me that our models of behavior and, more important, our recommendations for new policies and practices should be based on what people actually do rather than what they are supposed to be doing under the assumption that they are completely rational.”¹⁰

Clearly, in our agent-based models of complex social systems, we want to include agent behavior rules that reflect what people actually do. Our understanding of what real people actually do in social situations comes mainly from the work of Daniel Kahneman and Amos Tversky (see sidebar). Their experiments in human judgment and decision-making, together with the experiments of scientists such as Dan Ariely who followed them, uncovered the startling result that human judgment is based not on rational cognitive processes, but rather on *heuristics* – unconscious ‘rules of thumb’ that humans have developed over millennia to deal with our environment.

In his book *The origin of wealth*, Eric Beinhocker provides a vivid illustration of the distance between rationality and heuristics.¹¹ He asks you to imagine traveling on an airplane, sitting on the aisle next to an eccentric-looking woman.

Kahnemann and Tversky

Psychologists Daniel Kahneman and Amos Tversky shared one of the most productive collaborations in the history of social science. Starting in 1969, for more than 25 years they conducted groundbreaking experimental research into human judgment and decision-making. Their research had such a profound impact that in 2002 Kahneman became the first psychologist to win a Nobel Prize in Economics (an honor that, had he lived, Tversky would have shared).

As an example of one of their experiments:

In two trials, participants immerse a hand in cold water until instructed to remove it. The first trial lasts 60 seconds at 57 degrees Fahrenheit (very painful), and the second trial lasts a total of 90 seconds with 60 seconds again at 57 degrees followed by 30 seconds at 59 degrees (a little less painful). When asked which of the two trials they would choose to repeat, the remarkable finding is that 65-80 percent of subjects elect to repeat the second trial, even though it is longer than the first trial.

This, and a host of similar experiments, led Kahneman and Tversky to conclude that we store memories of our experiences according to what they called a “peak/end rule heuristic”: our memory of events is primarily an amalgam of the peak point of the experience and its end point. Nothing else matters.

For an excellent introduction to their work, see the YouTube videos of Kahneman presenting *Explorations of the mind*.⁹

⁹ Kahneman (2008a) and Kahneman (2008b)

¹⁰ Ariely (2008b), pages 279-281, one of this report’s *Top ten Complexity Science books*.

¹¹ Beinhocker (2006), pages 119-120, also one of this report’s *Top ten Complexity Science books*.

E. AGENT BEHAVIOR RULES CONTINUED

1. Behavior rule types continued

Heuristics continued

Mid way through the flight, the woman declares that she is bored, and says that she will give you and the businessman in the window seat \$1,000 if you can both agree how to share the money. She takes out ten \$100 bills to show that she is serious. The condition for her offer is that the businessman must decide how to split the money, and that you must accept his decision. If you reject his decision, you both get nothing. The businessman turns to you and says, “My decision is that I get \$900 and you get \$100. Would you accept his decision?”

Most people don't. This ‘ultimatum game’ has been played with thousands of real people of all types, from many cultures, and of all ages. The overwhelming majority reject the decision, because it is unfair. Clearly, the response is emotional and appears irrational, but recognizably human.¹²

Human behavior is marked by a slew of additional irrational heuristic biases with labels such as:

- framing bias
- anchoring bias
- representativeness bias
- availability bias
- confirmation bias
- conjunction bias
- narrative bias
- proximate cause bias
- expert bias

We also have great difficulty judging probabilities, and are supremely overconfident. No matter how much mathematics and statistics we study, or how much experience we have, we are human and we err.¹³

Physicians err, hospitals err, insurance companies err, consumers choosing health insurance policies err, corporate executives err, retirees err, all human and institutional agents in our Complexity Science models will err. We had best build this fact into our behavior rules.

¹² A case can also be made that in a larger cultural sense, the response is actually rational, for it fights unfairness.

¹³ To learn more about these biases and irrational behaviors, see Ariely (2008b), Kahneman, Slovic, & Tversky (1982), and Kahneman & Tversky (2000).

E. AGENT BEHAVIOR RULES CONTINUED

2. Developing behavior rules

There are many types of behavior rules and structures: simple if-then rules, genetic algorithm rules, neural network rules, game theory structures, behavioral economics rules, and many more that we have not covered. How do you decide which type to use?

In the foregoing discussion of rule types, I gave a few hints:

- For simple behavior, simple if-then rules or transition tables may be sufficient.
- Where agents need to search a large fitness landscape for optimal solutions, genetic algorithms may be helpful.
- Where agents need to learn from past patterns to determine future action, think of neural networks.
- Where cooperation or competition are involved, the structure of game theory may be appropriate.
- Whenever the agents are real people or institutions, consider the results of behavioral economics.

In practice, a combination of rule types may be appropriate.

You may find it helpful to:

- **Search the literature.** Perform formal literature searches for information about the way your model's agents behave and how to reduce that behavior to rules. For an example of a formal literature search, see the section at the end of this report titled *Finding the essential resources*.
- **Ask people.** Conduct interviews, focus groups, and surveys to elicit information about the ways humans and institutions behave.
- **Conduct controlled experiments.** Although outside the traditional province of actuaries, controlled behavioral experiments are one of the most powerful ways to understand human and institutional behavior. If you review the work of Ariely, Kahneman, and Twersky, you will discover that such experiments are easy to set up and administer.
- **Ask experts.** Use structured methods, such as 'knowledge engineering', to obtain information from experts. Knowledge engineering is a set of techniques from the field of software engineering for eliciting and organizing the knowledge of experts.¹⁴

¹⁴ North & Macal (2007), pages 103-112.

Complexity science – an introduction (and invitation) for actuaries

Chapter two: Agent-based models continued

F. COMPARISON WITH ACTUARIAL MODELS

Agent-based models and traditional actuarial models are quite different.

Although traditional actuarial models take many forms, in essence they simply project historical aggregate patterns into the future. These patterns may be population transition probabilities (such as rates of death, retirement, and disease); time series trend rates (such as interest rates or health expenditure trend rates); or probabilities of particular risks (such as rates of accident or catastrophe). And the projection may involve considerable actuarial judgment about deviations of future patterns from historical. But, no matter if the model type employed is a micro- (or cell-based) simulation, a statistical model, risk analysis model, or some other model type, the essential methodology is projecting historical aggregate patterns into the future, a methodology that is top-down.

By contrast, agent-based modeling is bottom-up. It seeks to understand and model the behavior of a system's fundamental units, its agents. System-wide attributes and behavior, such as the aggregate patterns of actuarial models, are then a by-product, an emergent result.

G. OBJECT-ORIENTED PROGRAMMING

Although agent-based models can be implemented using any programming language, an innovation in computer science called '**object-oriented programming**' fits naturally with agent-based models.

Object-oriented programming is used widely. It is the method used to create programs like Microsoft Excel and Word, most video games, and even movies like *Avatar*.

In object-oriented programming, any component of reality can be represented as an 'object'. For example, living things such as a person, a heart, and even a blood cell can be objects, as can inanimate things such as a phone and a photon.

Complexity science – an introduction (and invitation) for actuaries

Chapter two: Agent-based models continued

G. OBJECT-ORIENTED PROGRAMMING CONTINUED

In an agent-based model using object-oriented programming, each agent is an ‘object’ that is an instance of a ‘class’. For example, in an agent-based model of a healthcare system, a particular physician object named Dr. Welby might be an instance of the class called ‘physician’.

Every class has ‘attributes’ (variables) and ‘methods’ (functions). Each instance of a class will have a value for each attribute. For example, if ‘specialty’ is an attribute of the class ‘physician’, its value for Dr. Welby may be ‘primary care’, whereas its value for Dr. Jones may be ‘heart surgeon’. Attributes can change over time. For example, the value of the attribute ‘address’ may change for Dr. Welby when he moves from the city to the country.

Methods are the functions of a class. For example, a method for the class ‘physician’ might be a behavior rule to prescribe a certain medication when a patient presents with a specified constellation of signs and symptoms.

Agents are naturally implemented as instances of classes, with their attributes and state information corresponding to attributes, and their behavior rules corresponding to methods. A population of agents is easily created by generating multiple instances of a class. As you will see in Chapter six, the many agents of Sugarscape and Archimedes are generated as instances of object-oriented classes.

Just as there can be hierarchies of agents, there can be hierarchies of classes, such as the class ‘hospital’, the sub-class ‘employee’, the sub-sub-class ‘physician’, etc., down to any level of detail.

The attributes and methods of a class are ‘inherited’ by every instance of that class. For example, the class ‘physician’ may have ‘skill level’ as an attribute. This attribute would be inherited by the class instance Dr. Welby. Inheritance makes object-oriented models easy to update: Whenever an attribute or method is changed for a class, the change is automatically inherited by all instances of the class and its subclasses.

Many modern programming languages support object-oriented programming, including Java, VB.NET and C++.

H. STRENGTHS AND WEAKNESSES

Agent-based models are particularly advantageous when:

- A system can be represented by interacting agents with definable behaviors.
- Relationships among agents, or between agents and the environment, can change over time.
- Agent dynamics are dependent on spatial relationships, such as geographic location.
- Agents can adapt their behavior or learn.
- You need to present a model to people who are uncomfortable with mathematical models. People can often relate to the notion of agents more easily than to abstract mathematical models. They can easily imagine taking on the role of an agent.

They are not as useful when:

- A system is simple, with few component parts.
- Agent behavior cannot be modeled, even approximately.
- It is impossible to gather the data necessary to establish individual agent attributes.

I. GETTING STARTED

Perhaps the best way to get started with agent-based modeling is to experiment with and understand the models accompanying this report. Next, do the report's exercises; many of these ask you to modify the models accompanying this report. Then, find a problem in your work that appears to be amenable to an agent-based modeling approach, and build a model to address the problem.

As you build agent-based models, you will find it helpful to follow common modeling best practices (see sidebar).

Best practices

In their book *Complex adaptive systems*, John Miller and Scott Page list several best practices for agent-based modeling.¹⁵ Following are extracts:

Keep the model simple: Good models strip phenomena down to their essentials, yet retain sufficient complication to produce the needed insights.

Focus on the science, not the computer: New technological developments may enhance our ability to explore better existing models or create new ones, but without a solid model underlying the work, such improvements are meaningless.

Avoid black boxes: Each part of a model must be as clear and accessible as possible.

Nest your models: Nesting standard models within computational ones is usually a very natural process. ... once nested, it is easy to compare the model's predictions in the special cases with known results, and then to show how the model verifies known results and observations. One way to nest models is to rely on "tunable" dials for controlling key assumptions.

Create multiple implementations: One useful way to facilitate the creation of multiple implementations of a model is to have at least two groups separately code the model (preferably using two different computer languages). Not only does this process help clarify the important issues, but it also results in two versions of the model that can be run in parallel to confirm results and gain insights.

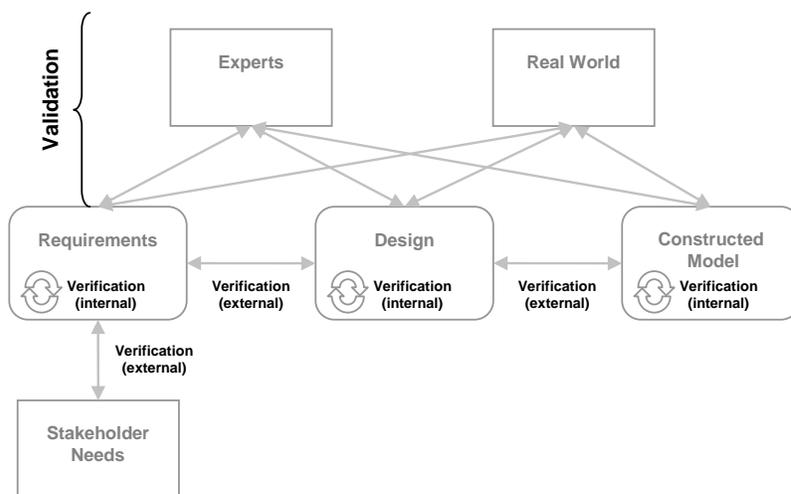
(continued on the next page)

¹⁵ Miller & Page (2007), pages 245-254.

J. VERIFICATION AND VALIDATION

One of the best practices recommended by Miller and Page is to ‘prove your results’ (see sidebar). The accepted way to prove your results is called ‘verification and validation’, or ‘V&V’.

Verification and validation are often confused with each other.¹⁷ As illustrated in the diagram below, model verification involves both externally directed tasks and internally directed tasks. Externally, verification ensures that the model is an accurate reflection of stakeholder needs, that design accurately follows requirements, and that construction accurately follows design. Internally, it ensures that the model is internally consistent and without defects.



Validation involves two externally directed tasks. One ensures that the model is an accurate reflection of the real world, and the other ensures that experts assess the model as reasonable, practicable, and relevant. Standard actuarial measures, such as the coefficient of determination (R^2), can be used to ensure that the model accurately reflects the real world.

It is best practice to include structured V&V processes, such as structured walkthroughs or formal audits, at each major step of model development such as after requirements analysis, design, and construction.¹⁸

Best practices continued

Check the parameters: ... computational models should always be subject to sensitivity analysis of key parameters.

Document code: Care and time spent in this domain are necessary for ensuring that the results can be fully analyzed and easily tied to the exact conditions that produced the outcome.

Beware of debugging bias: When modelers observe results that are not as expected, they are likely to spend a lot of effort debugging their code. When their expectations are met, little such effort is expended.

Write good code: McConnell (2004) provides a nice overview of the basics for writing high-quality, extendable, easily communicated code.

Distribute your code: Code for published results should be easily available to others so that they can replicate the results.

Prove your results: Whenever possible, computation results should be clarified and verified as thoroughly as possible ... Whenever possible, the analysis of computational models should be enhanced with complementary modeling efforts.

Reward the right things: ... judgments in this area should focus not on the computer per se, but on the quality and simplicity of the model, the cleverness of the experimental designs, and the new insights gained by the effort.¹⁶

¹⁶ Miller & Page (2007), pages 245-254.

¹⁷ They can also be confused with the epidemiologic concepts of internal and external validity, which relate to proper demonstration of cause-effect relationships between variables in scientific studies (internal validity) and to whether such relationships can be generalized (external validity).

¹⁸ For more information about V&V, read chapter 11 of North & Macal (2007).

Complexity science – an introduction (and invitation) for actuaries

Chapter two: Agent-based models continued

K. HISTORY

Although agent-based models were first contemplated in the 1940s, because of limited computing power, they were not implemented.

In the 1970s, Conway's Game of Life (see Chapter five) was one of the first computerized agent-based models. It was soon followed by Schelling's segregation model (see Chapter six) in which agents moved around on a two-dimensional grid.

In 1987 at the Santa Fe Institute, Craig Reynolds presented his simulation of flocking birds (that he called 'boids'). Each boid followed three simple in-flight rules and, with no other controller, the result was so lifelike that the approach has been used for many movies to simulate flocking and stampeding behavior.¹⁹ The Santa Fe Institute continued to support the development of agent-based models, and sponsored the first widely used agent-based modeling platform, SWARM.

The first large-scale agent-based model was Sugarscape, developed by Joshua Epstein and Robert Axtell at the Brookings Institution in 1996. In Chapter six, we will explore Sugarscape in detail.

In the 1990s, Uri Wilensky and others developed the popular agent-based educational platform NetLogo. And in 2000, Argonne National Laboratory released the agent-based modeling platform called Repast. In Chapter six, we will use Repast to develop the Schelling Segregation model and Sugarscape.

L. EXERCISES

1. Find a counter-example to the author's statement that traditional actuarial models project historical aggregate patterns into the future. (When you find it, please let the author know.)
2. Identify two areas in your work that can benefit from agent-based modeling.
3. One of Joshua Epstein's sixteen reasons for building models is to "discover new questions" (number five). Do you think Complexity Science has the potential to prompt you to ask questions about your work that you've never asked, or to explore areas you previously thought were impossible to address?
4. Go through the details of the ANN perceptron model in section E, to convince yourself that it works.

¹⁹ To see an example of 'boids' visit Craig Reynolds web page "www.red3d.com/cwr/boids/".

Complexity science – an introduction (and invitation) for actuaries

Chapter two: Agent-based models continued

M. TO LEARN MORE

To learn more about agent-based modeling, you may enjoy the book *Managing business complexity* by Michael North and Charles Macal.²⁰ You may also enjoy reading two chapters in Joshua Epstein's book *Generative social science* titled:

- *Agent-based computational models and generative social science*, and
- *Remarks on the foundations of agent-based generative social science*.²¹

To learn more about genetic algorithms, read chapters 9 and 11 of Melanie Mitchell's book *Complexity*.²² You may also enjoy her earlier book *An introduction to genetic algorithms*.²³

To learn more about artificial neural networks, read *Fundamentals of neural networks* by Laurene Fausett.²⁴ To learn more about the application of game theory to Complexity Science, read Robert Axelrod's book *The complexity of cooperation*.²⁵

To learn more about heuristics and behavioral economics, read Dan Ariely's book *Predictably irrational*.²⁶ You may also enjoy watching his *Predictably irrational* series on YouTube.com.²⁷ For information about behavioral economics and its applications (including applications to healthcare) read *Behavioral economics and its applications*.²⁸ For applications of behavioral economics to retirement issues, read *Behavioral dimensions of retirement economics*.²⁹

To learn more about building credible agent-based models, and about their verification and validation see the papers *How to build valid and credible simulation*³⁰ models and *Verification and validation of simulation models*.³¹

²⁰ North & Macal (2007), one of this report's *Top ten Complexity Science books*.

²¹ Epstein (2006), pages 1-74, also one of this report's *Top ten Complexity Science books*.

²² Mitchell (2009)

²³ Mitchell (1996)

²⁴ Fausett (1994)

²⁵ Axelrod (1997)

²⁶ Ariely (2008b), one of this report's *Top ten Complexity Science books*.

²⁷ Ariely (2008 - 2009)

²⁸ Diamond, Vartiainen, & Yrjö Jahnssonin säätiö. (2007)

²⁹ Aaron (1999)

³⁰ Law (2009)

³¹ Sargent (2009)

Complexity science – an introduction (and invitation) for actuaries

Chapter two: Agent-based models continued

N. REVIEW AND A LOOK AHEAD

This chapter concludes Part I. It introduced agent-based models and discussed their characteristics, their strengths and weaknesses, how they compare with actuarial models, their history, and how you can start using them.

In Part II, we will examine the four archetypal agent-based models of Complexity Science. After reading its chapters and doing the exercises, you will have a firm grasp of Complexity Science models and how to build them.

PART II: COMPLEXITY SCIENCE MODELS

Models can surprise us, make us curious, and lead to new questions. This is what I hate about exams. They only show that you can answer somebody else's question, when the most important thing is: Can you ask a new question? It's the new questions ... that produce huge advances, and models can help us discover them.

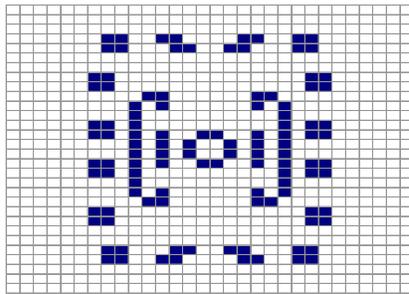
Joshua M. Epstein¹

¹ See J. M. Epstein (2008), section 1.15.

CHAPTER THREE: FOUR ARCHETYPAL MODELS

A. INTRODUCTION

In October 1970, an unusual article appeared in *Scientific American*, in the ‘Mathematical Games’ column. Its title was “The fantastic combinations of John Conway’s new solitaire game ‘life’” and was about British mathematician John Conway’s game called the Game of Life.¹ The article traced how Conway developed the game, gave the game’s simple rules, and reported some of its theoretical results. But that’s not what got people’s attention. Rather, people were, and continue to be, fascinated by the lifelike patterns generated by the Game of Life. As its computation unfolds, the game’s grid becomes a canvas of complex patterns that seem alive, a miniature universe that evolves (see the sidebar.)



The Game of Life was one of the first powerful examples of ‘cellular automata’, one of the four archetypal model types of Complexity Science that you will soon learn how to apply.

About the same time, Thomas Schelling, a prominent economist and recipient of the 2005 Nobel Prize in Economics, was on a plane flight from Chicago to Boston. He began thinking about racial segregation, and drawing Xs and Os on a grid, like tic-tac-toe. The Xs and Os represented people of two different races that move from space to space on the grid according to simple rules driven by their racial preferences. With this simple model, he discovered a startling fact: even color-blind preferences lead to neighborhoods that are segregated.³

There was something alive ...

In his book *Complexity*, Mitchell Waldrop presents Chris Langton describing his encounter with The Game of Life² (Langton is one of the pioneers of Complexity Science.):

“One time I glanced up,” he says. “There’s the Game of Life cranking away on the screen. Then I glanced back down at my computer code – and at the same time, the hairs on the back of my neck stood up. I sensed the presence of someone else in the room.”

Langton looked around, sure that one of his fellow programmers was sneaking up on him. ... But no – no one was behind him; no one was hiding. He was definitely alone.

Langton looked back at the computer screen. “I realized it must have been the Game of Life. There was something *alive* on that screen. And at that moment, in a way I couldn’t put into words at the time, I lost any distinction between the hardware and the process.”

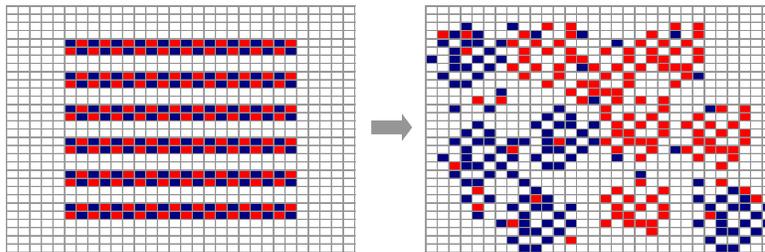
¹ See Gardner (1970). To watch Conway tell the story of the Game of Life, see Conway (2007a) and Conway (2007b). Rules for the Game of Life are given in Chapter five, and you can play the game on the Excel two-dimensional cellular automaton model that accompanies this report.

² Waldrop (1992), pages 202-203.

³ Schelling (2006)

A. INTRODUCTION CONTINUED

Schelling’s Segregation model was one of the first ‘artificial society’ models of Complexity Science, another of the four archetypal model types.



The Game of Life and Schelling’s Segregation Model are two early examples of the computer-based models and bottom-up perspective that distinguish Complexity Science from traditional science. In Part II, you will learn how to apply these new tools and perspectives to solve real-world problems.

B. FOUR ARCHETYPAL MODELS

Part II introduces you to four archetypal model types that span all Complexity Science models, and shows you how to use them. The following table shows the salient features of the four archetypal model types.

Model type	Agents	Agent relationships	Agent behavior rules	Environment	User involvement
1. Networks	●	●			
2. Cellular automata	●	●	●		
3. Artificial societies	●	●	●	●	
4. Serious games	●	●	●	●	●

B. FOUR ARCHETYPAL MODELS CONTINUED

1. Networks

The first archetype is networks. With networks, we will explore models of agents and their relationships. For example, in a complex real-world economic system, two people (agents) might be related because they are neighbors.

As you will see, one can learn a lot about a complex system by using networks to explore its underlying structure. The branch of Complexity Science that developed this model type is ‘**network science**’ or ‘**network theory**’. Chapter four introduces networks and shows you how to apply them.

2. Cellular automata

In the next type of archetypal model, cellular automata, agents and their relationships are augmented by agent behavior rules. In an economic system, for example, people (agents) who are neighbors (relationship) might engage in trade with one another (following agent behavior rules).

This model type expands our capacity to understand real-world systems. In Chapter five, you will learn about cellular automata and how to apply them.

3. Artificial societies

The next archetype, artificial societies, adds an environment. Now, in addition to relationships and interactions among themselves, agents also interact with an environment. For example, in an economic system, neighboring people might trade with one another, and might also extract products from an environment (such as fruit or gold) to trade.

Chapter six introduces artificial societies and shows you how to use them.

4. Serious games

The fourth archetype is serious games. These models incorporate one or more active users who play a game that usually incorporates agents, relationships, agent behaviors, and an environment. The primary purpose of such games is for the users to better understand the real-world system being modeled. You will learn about serious games in chapter seven.

Complexity science – an introduction (and invitation) for actuaries

Chapter three: Four archetypal models continued

B. FOUR ARCHETYPAL MODELS CONTINUED

Boundaries among these archetypal models are blurred. For example, network science also includes a type of network called a ‘**dynamical network**’ that incorporates agent behavior rules within networks. As we progress, I will point out significant places where the boundaries are blurred. Nevertheless, these four archetypes will help you understand and apply Complexity Science models.

C. STRUCTURE OF PART II

Each of the four chapters of Part II presents one of the model archetypes. Each includes:

- An introduction to the model type, including definitions and theoretical examples
- A discussion about how the model type relates to the other model types
- An introduction to the new concepts and perspectives of Complexity Science that underlie applications of the model type
- Examples of real-world applications
- Supporting material to help you learn to apply the model type (see sidebar)

Supporting material

This report’s supporting material will help you learn about Complexity Science:

Computer code

An enjoyable and effective way to learn about Complexity Science is to play with computer models. The computer code used to generate examples of the archetypal models in the next four chapters are available on the SOA web page for this report.

To make it easy to work with the models, the code includes many comments.

The network models in this report were generated using the modeling platforms ‘R’ with ‘igraph’, ‘Excel’ with ‘Visual basic for applications’, and ‘Java’ with ‘Repast Symphony’. To set up these platforms on your computer, see the document on the SOA web page titled *Getting started with modeling platforms*.

Exercises

To help you solidify and apply what you have learned, at the end of each chapter is a section with (I hope) entertaining and useful exercises.

References

Following the exercises are references to resources that will help you learn more.

Glossary

At the end of this report is a *Glossary* of Complexity Science terms introduced in the report.

CHAPTER FOUR: NETWORKS

Thanks to the rapid advances in network theory it appears that we are not far from the next major step: constructing a general theory of complexity. The pressure is enormous. In the twenty-first century, complexity is not a vague science buzzword any longer, but an equally pressing challenge for everything from the economy to cell biology. Yet, most earlier attempts to construct a theory of complexity have overlooked the deep link between it and networks. In most systems, complexity starts where networks turn nontrivial.

Albert-László Barabási¹

A. INTRODUCTION

About the time that Conway was publishing the Game of Life, and Schelling was modeling segregation with X's and O's on a grid, a young sociologist named Stanley Milgram was starting an experiment that would become famous as the “small world experiment”. He was curious about the structure of the social network that connects people. His intuition was that two random people in this network are closer together than one might expect. To test his intuition he mailed 160 packages to a random group of people in Omaha, Nebraska, asking each to forward the package to an acquaintance who might personally know a named stockbroker in Boston, and thus would be able to deliver the package. The next recipient was asked to do the same.

Given that the population of the U.S. at that time was about 200 million, how many links on average do you think it took between the 160 Nebraskans and the Boston stockbroker? People who don't know the answer usually guess one hundred or more. But you probably already know the answer. It was about six. Thus, on average, there were ‘six degrees of separation’ (see sidebar) between them. To test this idea further, Milgram sent packages to randomly selected whites in Los Angeles, asking them to get the packages to randomly selected blacks in New York. Again, this time even more surprisingly, the average number of links was about six. Indeed, given the results of other studies, it may well be that two random people among the earth's seven billion may be only six acquaintances distant.³ It appears to be a small world.

Small world

The first published appearance of the small world insight – that the world's inhabitants are connected by no more than six links – was not by a mathematician, but a celebrated Hungarian short-story writer named Frigyes Karinthy.

In his 1929 story titled *Chain-links*², he writes:

“Planet Earth has never been as *tiny* as it is now. It shrunk – relatively speaking of course – due to the quickening pulse of both physical and verbal communication. ... One of us suggested performing the following experiment to prove that the population of the Earth is closer together now than they have ever been before. We should select any person from the 1.5 billion inhabitants of the Earth – anyone, anywhere at all. He bet us that, using no more than *five* individuals, one of whom is a personal acquaintance, he could contact the selected individual using nothing except the network of personal acquaintances.”

Nearly forty years would pass before Milgram would start his experiments, and another thirty years before we would understand why our planet is a small world.

¹ See Barabási (2003), page 237. This is one of this report's *Top ten Complexity Science books*.

² To read the short story, see M.E.J. Newman, Barabási, & Watts (2006), pages 21-26, also one of this report's *Top ten Complexity Science books*.

³ For example, a Microsoft 2008 study titled *Planetary-scale views on an instant-messaging network* shows that the average chain of contacts between users of its worldwide .NET Instant Messenger Service was 6.6 people.

INTRODUCTION CONTINUED

But how is the so-called ‘**small world effect**’ possible? If I were to give you a sheet of paper with seven billion dots on it representing the earth’s population, how could you connect the dots such that there are only six links between any two dots, and yet retain the characteristic of most human communities that most connections are among dots close together? ... It took thirty years to find the answer to this question. Using the simplest model type in Complexity Science, the network, you will soon learn the answer.

A ‘**network**’ is a collection of real-world entities (agents) connected by relationships. Examples of networks are social systems of people related by friendship or consanguinity, an organization’s employees related by an organizational chart, and healthcare providers related by referral relationships.

The goal of network models is to understand how the structure of networks affects the behavior of complex systems built upon them. We would like to understand, for instance, how the structure of social networks affects the spread of disease, how the structure of an organization affects its vulnerability, and how the structure of a provider community affects access to health care, quality, and expense.

The modern study of networks is still in its infancy. It started around the turn of this century, with the publication of two short but revolutionary papers, one by Watts and Strogatz titled “Collective dynamics of ‘small-world’ networks”, and one by Barabási and Albert titled “Emergence of scaling in random networks”, papers that we will explore in this chapter.⁴

The best way to learn about networks is to use them. This chapter provides many tools and exercises to help you learn. In particular, don’t hesitate to learn the statistical software package ‘R’ that this chapter introduces; it is powerful and easy to use, as is its component package for network analysis called ‘igraph’. Setup instructions for ‘R’ and ‘igraph’ are in the separate document *Getting started with modeling platforms*.

⁴ Both articles are in M.E.J. Newman, et al. (2006), (one of this report’s *Top ten Complexity Science books*). The Watts-Strogatz paper is on pages 301-303, and the Barabási-Albert paper is on pages 349-352.

B. NETWORK BASICS

Now, let's cover basic network model definitions and concepts.

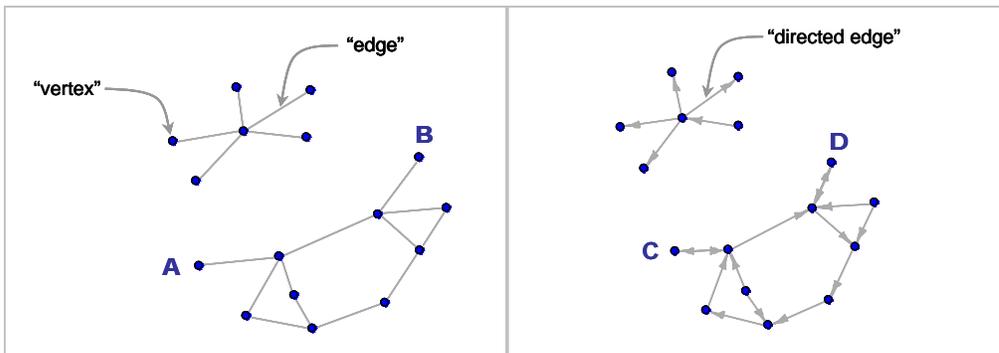
A **'graph'** is a representation of a real-world network. It consists of a collection of points called **'vertices'** that are connected by lines called **'edges'** (see the two simple graphs below⁵). A vertex represents a network agent and an edge represents a relationship between two agents (see the sidebar).

Vertices and edges can be either homogeneous (all alike) or heterogeneous (different). Vertices can also be hierarchical (ie, a vertex can represent a whole network). An edge can be either **'directed'** (representing a directional relationship between two nodes, such as an email sent from one person to another), or **'undirected'** (the relationship is bi-directional). A graph with directed edges is called a **'digraph'**.

A vertex is a node is a ...

Because Complexity Science is a new field that straddles many traditional fields, its terminology can become a little confusing.

For example, the terminology for graphs can come from mathematics, sociology, computer science, or physics. What we will call a vertex can also be called a node (computer science), an actor (sociology), or a site (physics). Similarly, an edge is also called a link (computer science), a tie (sociology), and a bond (physics). In the body of this report, I present the most commonly used terms in Complexity Science, but include alternative terms in the Glossary.



The **'geodesic path'**, or simply **'geodesic'**, between two vertices is the path with the least number of edges that must be traversed to travel from one vertex to the other. In the undirected graph on the left above, the geodesic path length from A to B (or B to A) is 3. In the directed graph, what is the geodesic path length from C to D? And from D to C? (Hint: they are different.)

The **'mean geodesic'** is the average of all a graph's geodesics. For the undirected graph above, the mean geodesic is 2.02.

⁵ Created with the function CreateSimpleGraph() and plotted with PlotGraph(), both from the set of igraph functions supplied with this report. For example, to create the undirected graph: in R enter gGraph1 <- CreateSimpleGraph(FALSE) followed by PlotGraph(gGraph1,"FR", FALSE, 7, "Console", "", 2,2).

B. NETWORK BASICS CONTINUED

The ‘**diameter**’ of a graph is the number of edges of the longest geodesic path. The diameter of the undirected graph on the previous page is 4. Can you see why? What is the diameter of the directed graph?⁶

The ‘**degree**’ of a vertex is the number of edges connected to it. A vertex of a directed graph has an ‘**in-degree**’ (the number of in-coming edges) and an ‘**out-degree**’ (the number of out-going edges). The ‘**maximum degree**’ is the degree of the vertex with the greatest degree. For the undirected graph on the previous page, the maximum degree is 5. Likewise, the ‘**minimum degree**’ is the degree of the vertex with the least degree (1 for the undirected graph). The ‘**mean degree**’ is the average of the degrees for all vertices (2.13 for the undirected graph).

A vertex’s neighboring vertices within a certain edge distance are called a ‘**neighborhood**’. Vertex ‘A’ on the simple undirected graph has four neighbors within its neighborhood of two edges.

Each of the simple graphs on the previous page consists of two groups of vertices, like islands separate from one another. Each such island is a ‘**cluster**’. So, the simple graphs have a cluster number of 2.

The ‘**density**’ of a graph is the ratio of its actual number of edges to the possible number of edges (0.133 for the undirected simple graph). Its ‘**transitivity**’ is a measure of the probability that the adjacent vertices of a vertex are connected. It is equal to:

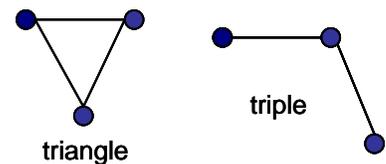
$$\frac{3 \times \text{number of triangles in a graph}}{\text{number of connected triples}}$$

A ‘**triangle**’ is three vertices all of which are connected to each other, and a ‘**connected triple**’ is three connected vertices (which may also be a triangle). For the undirected graph, the number of triangles is 1, and the number of triples is 32, so the graph’s transitivity is 3/32, or 0.09375. Transitivity measures the density of triangles in a graph, or its degree of cliquishness.

Why so many definitions?

Why do we need so many definitions and measures, when we can simply look at a graph to understand its characteristics and structure?

The main reason is that Complexity Science deals with the real world, and real-world networks are often big. For a network with scores or even hundreds of agents, the structure of its graph is easily captured by the eye. But the networks of Complexity Science may have millions or billions of agents, and their graphs cannot be effectively rendered for viewing even with modern computer tools. To understand these networks, we need careful definitions and statistical measures.

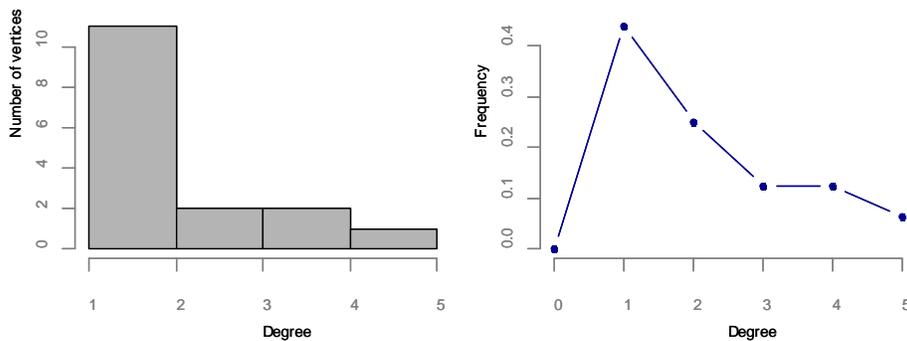


⁶ It is 7.

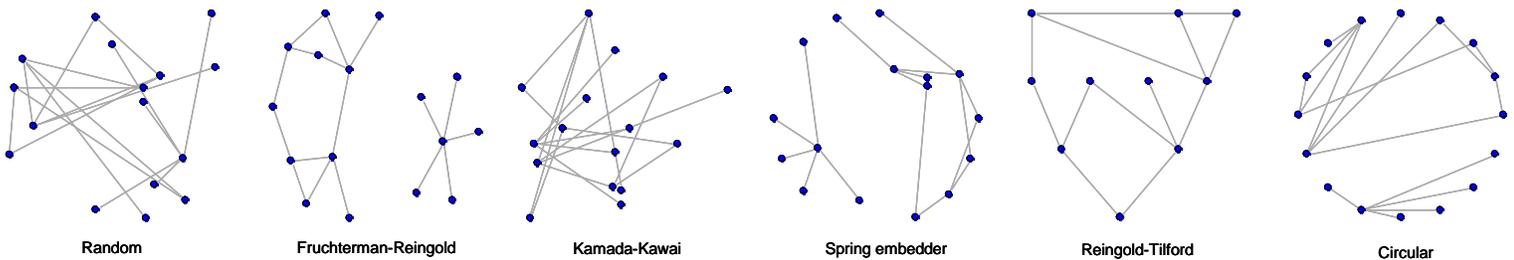
B. NETWORK BASICS CONTINUED

A graph’s ‘**correlation coefficient**’ measures the extent to which vertices are connected to other vertices with like degree (i.e., large-degree vertices connected to other large-degree vertices, and small-degree connected to small-degree).⁷ It varies between -1 (negatively correlated) to 1 (perfectly correlated). The correlation coefficient of the simple undirected graph is -.83. Can you see why?

The following charts show the ‘**degree distribution**’ of the simple undirected graph above, presented both as a histogram and as a frequency distribution.⁸ Can you verify them?



Graphs can be displayed in various ‘**layouts**’, which are alternative ways to place graph vertices. For example, below are six different layouts for displaying the simple undirected graph we have been discussing. The circular layout places the vertices equidistant on a circle, and the Reingold-Tilford layout places one of the clusters in a hierarchy. Above I used the Fruchterman-Reingold layout to display the simple graph, because it displays the graph’s structure clearly. For more information about the layouts, refer to the ‘layout’ section of the igraph documentation.⁹



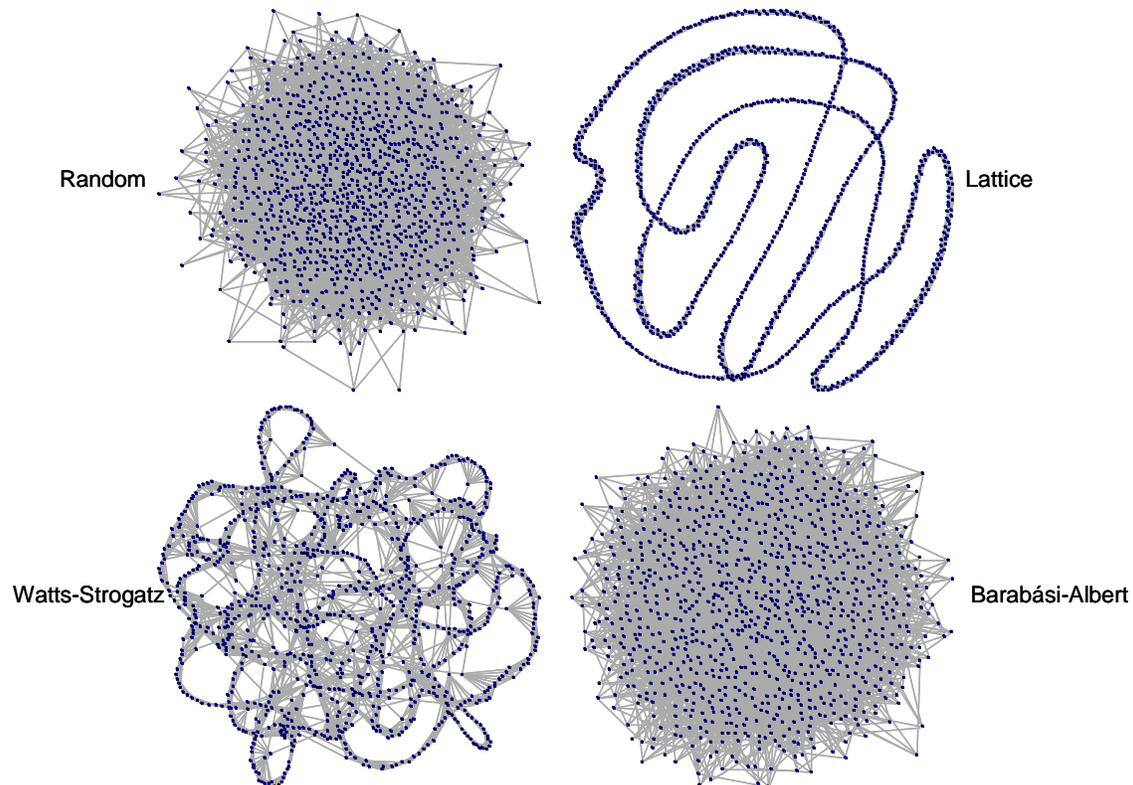
⁷ The tendency of vertices to connect to vertices with similar properties is called ‘assortative mixing’ or ‘assortativity’.

⁸ The charts were plotted with the functions PlotHistogram() and PlotDegreeDistribution().

⁹ The igraph documentation, titled “Package ‘igraph’” and written by Gabor Csardi, is available at “igraph.sourceforge.net”.

C. NETWORK MODELS

Now let's explore four primary models of networks in Complexity Science: the 'random', 'lattice', 'Watts-Strogatz', and 'Barabási-Albert' models. Examples of these models, each with 1000 vertices, are shown below in the Fruchterman-Reingold layout:¹⁰



What do these graphs look like to you? The first looks, well, random: its vertices and edges do not seem to follow any pattern. The lattice model is a contrail of vertices. In the Watts-Strogatz model, most of the vertices are connected in long strings like Christmas lights. And the Barabási-Albert model looks like the random model (but, as we shall see, here looks are deceiving).

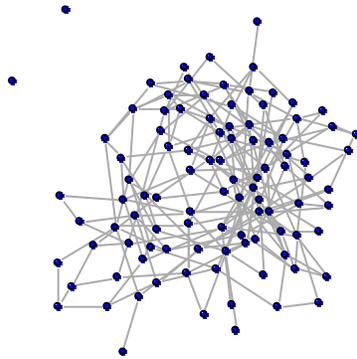
¹⁰ Created with `CreateRandomGraph(1000, .01, FALSE, FALSE)`, `CreateLatticegraph(1000, 5, FALSE)`, `CreateWattsStrogatzGraph(1000, 5, .01)`, and `CreateBarabasiAlbertGraph(1000, 5)`; plotted with `PlotGraph()` using the "FR" layout.

C. NETWORK MODELS CONTINUED

Let's see how each of these models is formed:

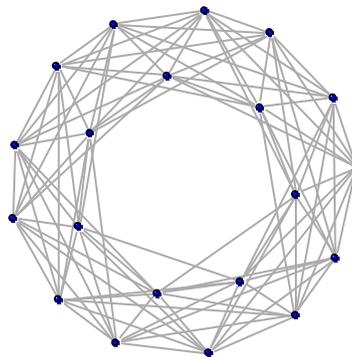
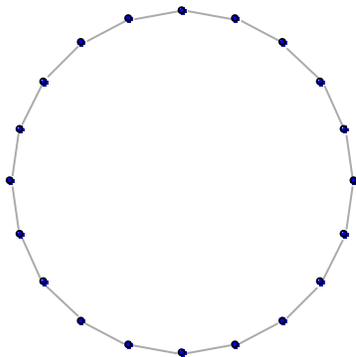
1. Random model

In the random network model (also known as the Erdős-Rényi model – see sidebar), the probability that two vertices are connected by an edge is fixed and independent of the vertices' attributes. For example, in the random graph of 1000 vertices above, the probability that two vertices are connected by an edge is 0.01. Following is a random network with an attachment probability of 0.01, but only 100 vertices.¹¹ Its connections still look random. Two of its vertices are even completely separated from the rest, like islands. (So, how many clusters does this graph have?)



2. Lattice model

In a lattice model, each vertex is connected to a fixed number ($2n$) of neighbors (n neighbors are on each 'side' of the vertex). In the lattice graph above with 1000 vertices, $n = 5$. Following are smaller versions, with only 20 vertices. The one on the left has $n=1$; the other has $n = 5$.¹²



Paul Erdős

The Hungarian Paul Erdős was one of the world's greatest mathematicians. He died in 1996 at age 83, and never had a wife, child, or property. Although he lived like an itinerate hobo with little to accompany him but a small suitcase, while traversing the world he authored more than 1,500 mathematical research papers, a few of which (co-authored with fellow mathematician Alfréd Rényi) were about the random network model.

Until recently, real-world networks of complex topology have been modeled using the random graph theory of Erdős and Rényi. But these models were rarely empirically tested. As you shall see, they are not the best models of reality.

A whimsical measure of relative stature among mathematicians is the 'Erdős number'. It is the geodesic between a mathematician and Paul Erdős in the network of all mathematicians, where an edge represents collaboration on a mathematical paper.

¹¹ Created with CreateRandomGraph(100, .01, FALSE, FALSE)

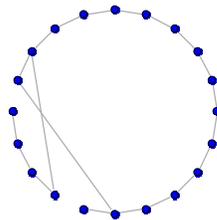
¹² Created with CreateLatticeGraph(20, 1, FALSE) and CreateLatticeGraph(20, 5, FALSE)

C. NETWORK MODELS CONTINUED

3. Watts-Strogatz model

In the Watts-Strogatz model, all of the vertices start out in a lattice network, connected like a string of pearls. Then, with probability p , an edge between vertex A and vertex B is replaced by an edge between vertex A and a randomly selected second vertex.

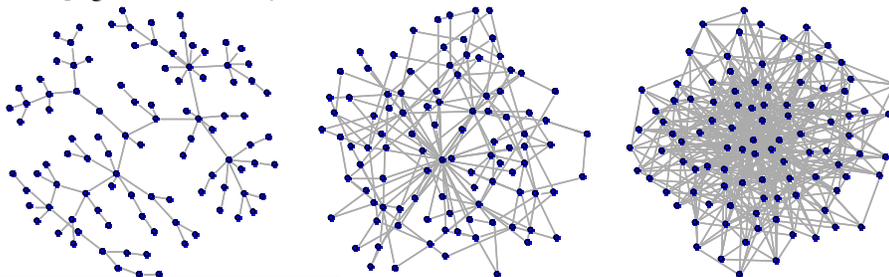
In the Watts-Strogatz graph above with 1000 vertices, each vertex in the starting lattice network has 2×5 neighbors, and the edge replacement probability is 0.01. Below is a starting lattice with 20 vertices, a 2×1 neighborhood, and a replacement probability of 0.10.¹⁴ Watts and Strogatz used such a graph to solve the small world problem (see sidebar).



4. Barabási-Albert model

The Barabási-Albert model starts with one vertex, and then at each subsequent step adds an additional vertex with m edges connected to previous vertices with a probability proportional to the degree of the vertices; thus, vertices with a higher degree are relatively more likely to become connected to additional vertices – the rich get richer.

For the Barabási-Albert graph above with 1000 vertices, the process took 999 steps, with five new edges being added at each step. Below are more simple Barabási-Albert graphs, with 100 vertices; the first adds one new edge at each step, the second adds two edges, and the third adds five edges.¹⁵ The second graph has the same number of vertices and edges as the random graph on the previous page. Now do they look different?



How Watts and Strogatz cracked the small world problem

As noted in the introduction of this chapter, it took about 30 years to solve the riddle of how two vertices in many real-world social networks (that are characterized by high local transitivity) can be separated by only about six degrees (six edges).

In their 1998 paper¹³, Duncan Watts and Steve Strogatz showed that real-world networks have high transitivity (their agents form cliques) and a low mean geodesic (low degrees of separation), two characteristics that cannot be modeled with either random graphs or lattice graphs (the two network models known at the time). They proposed a third model for the small world phenomena, and named it the small world model (we are calling it the Watts-Strogatz model).

With an edge replacement probability of only 0.01, this model has both high transitivity and a low mean geodesic. Thus, for a network to exhibit the small world property, they showed that it takes relatively few agents with connections outside of a clique (only about one in a hundred). As you will soon see, this small probability has great practical implications.

¹³ See M.E.J. Newman, et al. (2006), pages 301-303.

¹⁴ Created with CreateWattsStrogatzGraph(20, 1, 0.10)

¹⁵ Created with CreateBarabasiAlbertGraph(100, x) with $x = 1, 2, \text{ and } 5$.

C. NETWORK MODELS CONTINUED

Beyond the visual, how do these models compare? The following table statistically compares the examples with 1000 vertices shown above:¹⁶

	Random graph	Lattice graph	Watts-Strogatz graph	Barabasi-Albert graph
Number of vertices	1000	1000	1000	1000
Number of edges	4963	5000	5000	4995
Diameter	5	100	13	5
Mean degree	9.9	10	10	9.9
Maximum degree	21	10	12	185
Minimum degree	2	10	8	5
Mean geodesic	3.3	50.5	6.9	3.0
Density	0.01	0.01	0.01	0.01
Transitivity	0.01	0.67	0.63	0.03
Correlation coefficient	-0.009	NA	-0.012	-0.012
Number of clusters	1	1	1	1

In many respects the graphs are quite similar. But, a few of their measures indicate dramatic differences among them: They have markedly different diameters, maximum degrees, mean geodesics, and transitivity.

As the table shows, the mean geodesic of the lattice graph is much greater than any of the other three. It represents a ‘large world’ network, a type that is not commonly found in the natural or human social world. As a consequence, the diameter of the lattice graph is also greater.

Because the random graph has low transitivity, and thus does not correspond to the clique-ishness of the real world, it too does not qualify as a small world.

The Watts-Strogatz and Barabási-Albert graphs, by contrast, have low mean geodesics and higher transivities than the random graph. Thus, they are both potential models of small worlds.

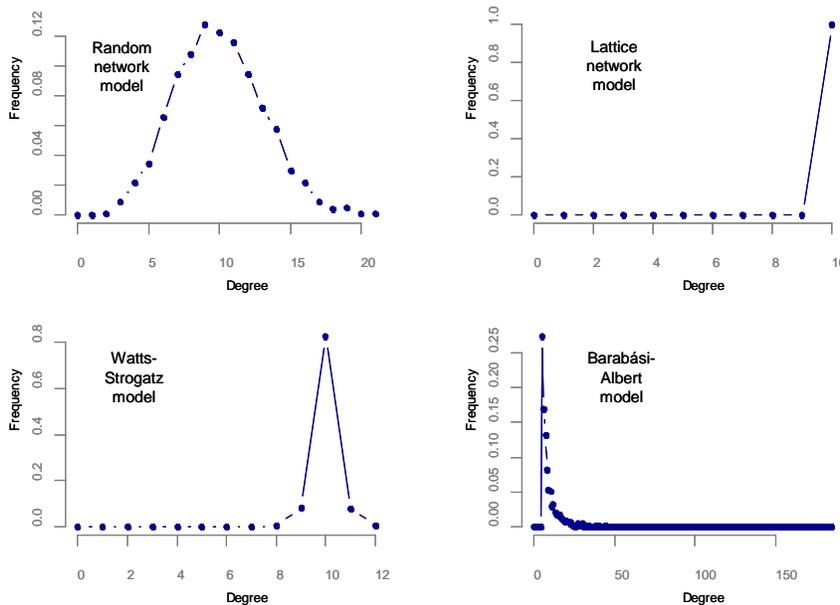
¹⁶ Developed using the function AnalyzeGraph().

C. NETWORK MODELS CONTINUED

The maximum degrees of the graphs are also quite different. Whereas the maximum degree of random, lattice, and Watts-Strogatz graphs are relatively small, the Barabási-Albert graph has a maximum degree of 185. This means that at least one vertex has a connection with 185 other vertices. A vertex with such a relatively large degree is called a ‘**hub**’, another characteristic of many real world networks. But it was yet another key feature of many real-world networks – the ‘**power law**’ – that prompted Barabási and Albert to develop their model.

In their famous 1999 paper¹⁹, they first showed that the degree distributions of many large real-world networks follow a power law (see sidebar). They then proposed a new model they called the ‘preferential attachment model’ (which we are calling the Barabási-Albert model) to represent such networks. The key factors of this model are (a) that it grows over time and (b) new edges are attached to vertices with a probability that is proportional to their number of degrees. Incorporating these two factors in a model results in a degree distribution that follows a power law.

Following are the degree distributions of the four model types:



The power law and Pareto

The degree distribution of many natural and human-made networks – including the World Wide Web (the global collection of web pages), the Internet (the interconnected underlying hardware), yeast protein interactions, and metabolic networks – follows a power law.¹⁷

The power law distribution formula is:

$$p(X = x) \sim x^{-k}$$

where k is a constant.

Or, the probability that a random variable takes the value ‘ x ’ is proportional to one over x to the k th power. For phenomena following the power law, small occurrences are common, whereas large occurrences are rare.

The power law is mathematically equivalent to the Pareto distribution and to ‘Zipf’s law’. Depending on the situation, one of these alternate forms may be more useful for analysis than the power law.¹⁸ For this report, we will typically use the Pareto distribution, rather than the power law distribution.

The Pareto distribution formula is:

$$p(X \geq x) = (x/x_{min})^{-k+1}$$

where x_{min} ($\neq 0$) is the minimum value of x .

This implies a linear relationship that is useful for identifying power law distributions:

$$\ln(y) = (-k+1)*\ln(x) - (-k+1)*\ln(x_{min})$$

¹⁷ See Clauset, Shalizi, & Newman (2007), pp 22-30 for an analysis of power laws in real data. This paper also presents a rigorous method to determine if a data set follows a power law.

¹⁸ See Adamic (2002).

¹⁹ See Barabasi & Albert (1999). This paper is found in M.E.J. Newman, et al. (2006), pages 349-352.

C. NETWORK MODELS CONTINUED

Clearly, the distributions are quite different. The degree distribution for the random graph looks like a Poisson distribution, which is not surprising, since the probability of a vertex having degree d is:

$$P(X=d) = \frac{n!}{d!(n-d)!} p^d (1-p)^{n-d}$$

This produces a binomial distribution, which, as n approaches infinity, becomes a Poisson distribution.

The lattice and Watts-Strogatz distributions are also straightforward: they are centered on degree 10, with a slight deviation in the Watts-Strogatz model due to its rewiring.

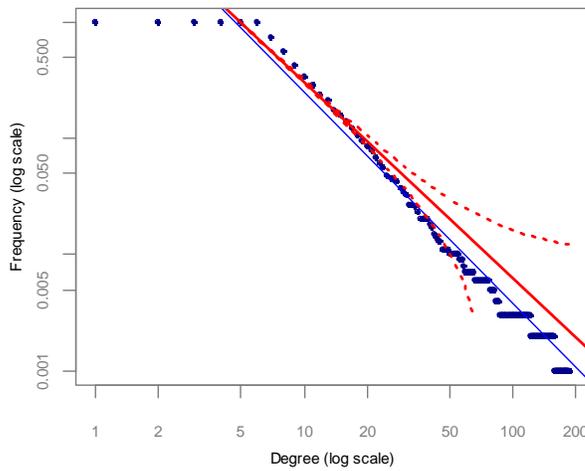
The Barabási-Albert model distribution is more unusual. It has many vertices with a low degree, but also has a long tail including a few vertices with very high degree. It is an example of a **‘skewed distribution’**, a non-symmetrical distribution with a long and often ‘fat’ tail.

Its many hubs are the ‘connectors’ mentioned in Malcolm Gladwell’s book *The tipping point*²⁰, and imbue these networks with special properties that we’ll study in this chapter. In the real social world, connectors are people who have connections with many other people.

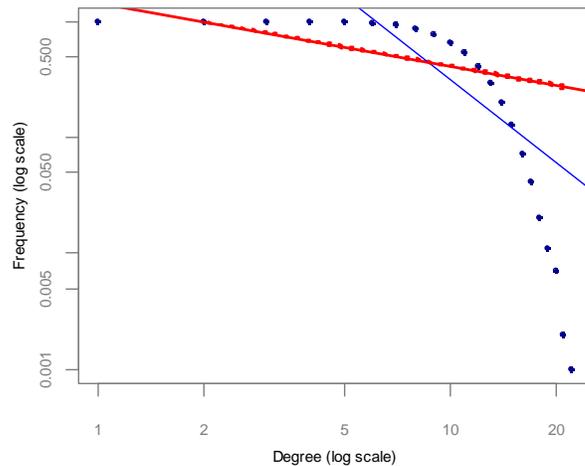
²⁰ Gladwell (2002)

C. NETWORK MODELS CONTINUED

The following chart shows that the degree distribution of our Barabási-Albert graph with 1000 vertices is nearly linear when plotted on log-log axes.²¹ In the chart, the solid red (thick) line is an exact Pareto (power law) distribution, and the dotted red lines are the Pareto distribution ± 0.01 . The solid blue (thin) line is the least squares regression line for the data points. Because the data points are close to the Pareto distribution, the degree distribution may follow a power law. (However, to conclusively prove this, much more work needs to be done.²²)



For comparison, following is a Pareto comparison chart for our random graph with 1000 vertices. Clearly, its degree distribution does not follow a power law (and neither do the degree distributions for the lattice and Watts-Strogatz graphs).



²¹ Plotted with PlotParetoComparison(gPAGraph, TRUE, 0, “Console”, “”, 5, 4).

²² For a rigorous method to determine if a degree distribution follows a power law, see Clauset, et al. (2007).

D. REAL-WORLD NETWORKS

Now that we have studied the basics of networks, as well as four key network models, let’s look at networks in the real world. The table below lists some of the human-made and natural networks that have been studied, together with their key features.²³

	Mean geodesic		Transitivity		Random	Small world	Power law
	Network	Random graph	Network	Random graph			
Human-made networks							
student romances	16.01	10.23	0.005	0.004	●		
company directors	4.60	3.64	0.59	0.002		●	
electric power grid	18.99	8.50	0.10	0.0006		●	
mathematics co-authorship	3.92	NA	0.15	0.00002		●	
email address books	5.22	5.33	0.17	0.001		●	
thesaurus word classes	4.87	3.28	0.13	0.009		●	
World Wide Web	11.27	NA	0.11	0.0002		●	●
Internet	3.31	5.40	0.035	0.0007		●	●
electronic circuits	11.05	NA	0.01	NA		●	●
Natural networks							
marine food web	2.05	2.46	0.16	0.06	●		
freshwater food web	1.90	1.76	0.20	0.240	●		
neural network	3.97	2.38	0.18	0.05		●	
metabolic network	2.56	3.17	0.09	0.012		●	●
yeast protein interactions	6.80	9.29	0.07	0.0006		●	●

The first and third columns of the table give the mean geodesic and transitivity for each of the networks, while for comparison the second and fourth columns give the mean geodesic and transitivity (where available) for random networks with the same number of vertices and edges as the real-world networks. For example, the mean geodesic for the network of company directors is 4.60 and the transitivity is 0.59, compared to 3.64 and 0.002 for a comparable random network. Thus, by definition, the network of company directors is a small world, as shown in the table.

As the table shows, the networks listed have three major types of structure: random, small world, and small world plus power-law distributed.

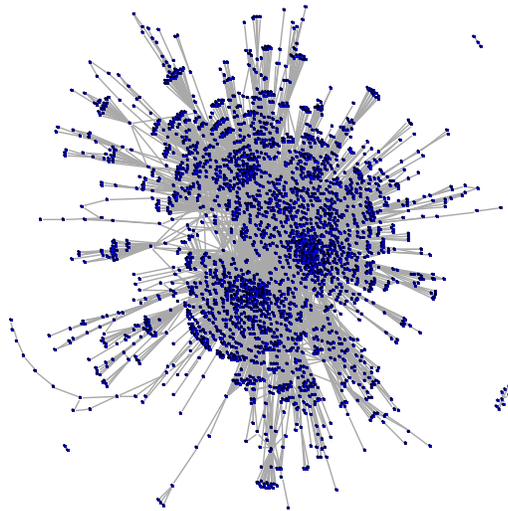
²³ See M.E.J. Newman (2003) pp 4-10 for a discussion of these networks, with citations. The mean geodesics and transivities for random graphs were calculated using the functions CreateRandomGraph() and AnalyzeGraph().

D. REAL-WORLD NETWORKS CONTINUED

Now let's explore two real-world networks in detail.

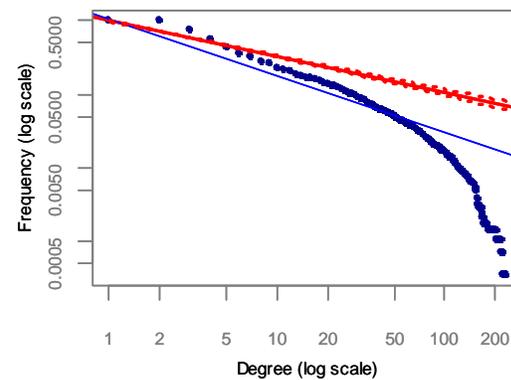
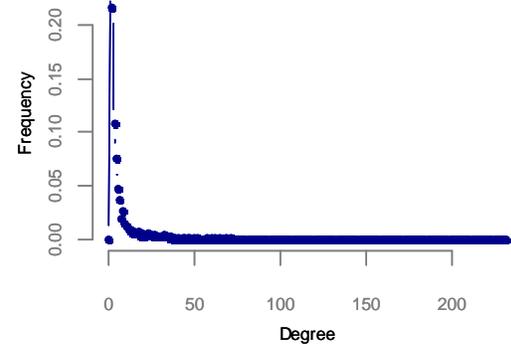
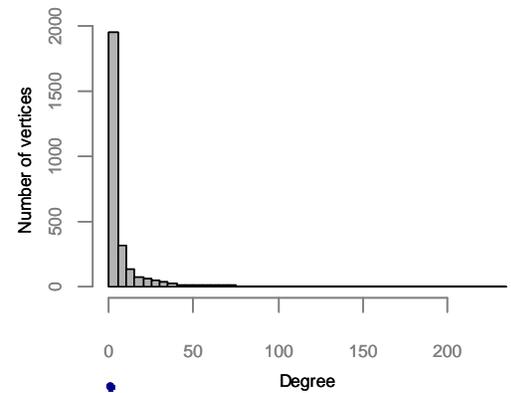
1. Airline routes

During one of your flights, while reading an in-flight magazine you have probably seen a map showing the airline carrier's routes from airport to airport. Below is a graph of all worldwide passenger flight routes, in the Fruchterman-Reingold layout.²⁴ It has 2,852 nodes (airports) and 14,996 edges (routes between airports).²⁵



Using its plot (above), and its key measures, density distributions, and Pareto comparison chart (below), what can you conclude about the real-world airline network?

Number of vertices	2,852
Number of edges	14,996
Diameter	16
Mean degree	10.5
Maximum degree	232
Minimum degree	1
Mean geodesic	3.95 (random = 3.64)
Density	0.004
Transitivity	0.25 (random = .004)
Correlation coefficient	-0.015
Number of clusters	4



²⁴ Created with CreateAirlineGraph() and PlotGraph().

²⁵ Determined with AnalyzeGraph().

D. REAL-WORLD NETWORKS CONTINUED

In 2000, there was a worldwide airline crisis. Around the world, because of demand exceeding capacity, airports were cancelling and delaying flights at an unprecedented rate (see the sidebar).

What led to such a worldwide crisis? And what could be done to alleviate it? To both questions, the network model type of Complexity Science provides an answer.

A clue to answer the first question is found in the Pareto comparison chart on the previous page. It – as well as the histogram and the density distribution chart – indicates that the degree distribution of airports nearly follows a power law, but falls away from a perfect power law at the higher-degree vertices. This suggests that airports grew into their current structure by preferential attachment, which indeed they did: for competitive advantage, airlines sought to link up with major airline hubs such as Atlanta, Chicago, and Los Angeles. The rich got richer. But, because available land for airport expansion is limited, there is a physical limit to an airport's size. For example, O'Hare and LAX cannot easily expand to meet growing demand. This limitation leads to the falling away from a perfect power law distribution at the higher degree vertices. Thus, the answer to the first question (what led to the worldwide crisis) is: the airline network evolved a preferential attachment structure in which high-degree hubs, because of the physical limitation of available land, could not continue growing to meet demand. Thus, the airline network's structure makes it susceptible to failure because of physical limitations. (The structure also makes it susceptible to terrorist attack, which we'll discuss below.)

And the answer to the second question? The point of an airline network is to enable a traveler to get from point A to point B in the shortest amount of time for the least cost. Translated to network terminology, such a network needs to be a small world. But, as you have seen, both the Watts-Strogatz model and the Barabasi-Albert models are small worlds. So, the answer to the second question is: transform the airline network to a network more like the Watts-Strogatz small world; encourage the modest expansion of smaller local airports with the inclusion of some long-distance flights for each. This is currently happening: smaller airports are starting to compete with the large hubs.

The airline crisis of 2000

In the spring of 2001, the U.S. Congress held a series of hearings to understand why the number of flight delays exceeding 45 minutes had doubled over the preceding five years, and why serious delays were affecting 10 percent of all flights. In the summer of 2000, O'Hare airport had cancelled some 4,600 flights and delayed an additional 57,000.

George Donohue, a professor of systems engineering and operations research at George Mason University, reported to Congress, "The U.S. hub and spoke air transportation system is approaching a serious capacity crisis." He pointed out that the network as a whole was then at 58 percent of capacity, and would soon be at 70 percent of capacity. The three largest airports, in Atlanta, Chicago, and Los Angeles, were at 80 percent of capacity. For all airports, serious delays start at about 50 percent of capacity.

The situation was the same around the world. Rod Eddington, chief executive of British Airways, said at the time, "Congestion in the skies and on the ground in Britain is becoming critical ... Heathrow is groaning at the seams."²⁶

Of course, September 11, 2001 changed the crisis from one of excess demand to one of inadequate demand. But, in time, the excess demand problem is sure to return.

²⁶ Much of this story comes from Buchanan (2002).

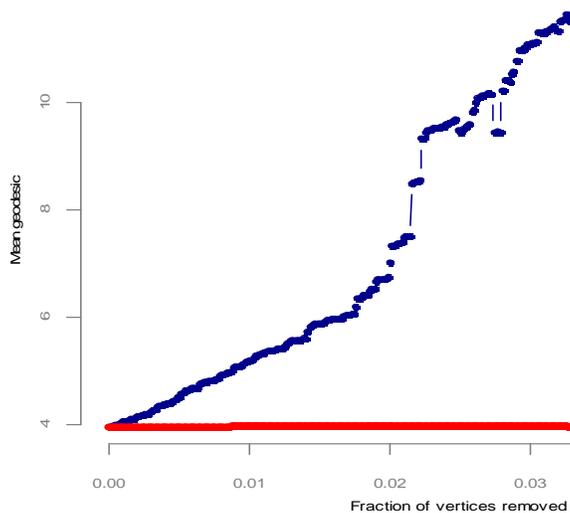
D. REAL-WORLD NETWORKS CONTINUED

Let's stay with the world airline network example for a moment more, to explore a few concepts of Complexity Science: 'emergence', 'resilience', 'fragility', and 'robustness'.

A key insight of Complexity Science is that complex systems 'emerge'; they are not engineered. Think of the difference between an airplane and the airline network. An airplane is a system, consisting of thousands of parts, that has been meticulously designed, constructed, and assembled for a specific purpose. Blueprints show exactly where each part goes, and manuals explain how the parts should be manipulated to achieve the airplane's purpose. An airplane is an extremely complicated system, but not a complex one.

By contrast, an airline network also consists of thousands of parts (airports) that serve a specific purpose, but its parts have grown up without a specific design. The network is a self-organized complex system that has emerged. Emergence is a key feature of complex systems that we will revisit many times in this report.

The airline network is simultaneously robust and fragile. The chart to the right shows what happens to the mean geodesic of the airline network as vertices are removed one by one.²⁷ The red horizontal line shows the result of removing vertices randomly, while the blue rising line shows the result of removing vertices in the order of their degree, starting with the highest-degree vertices. Do you see anything interesting about the chart?



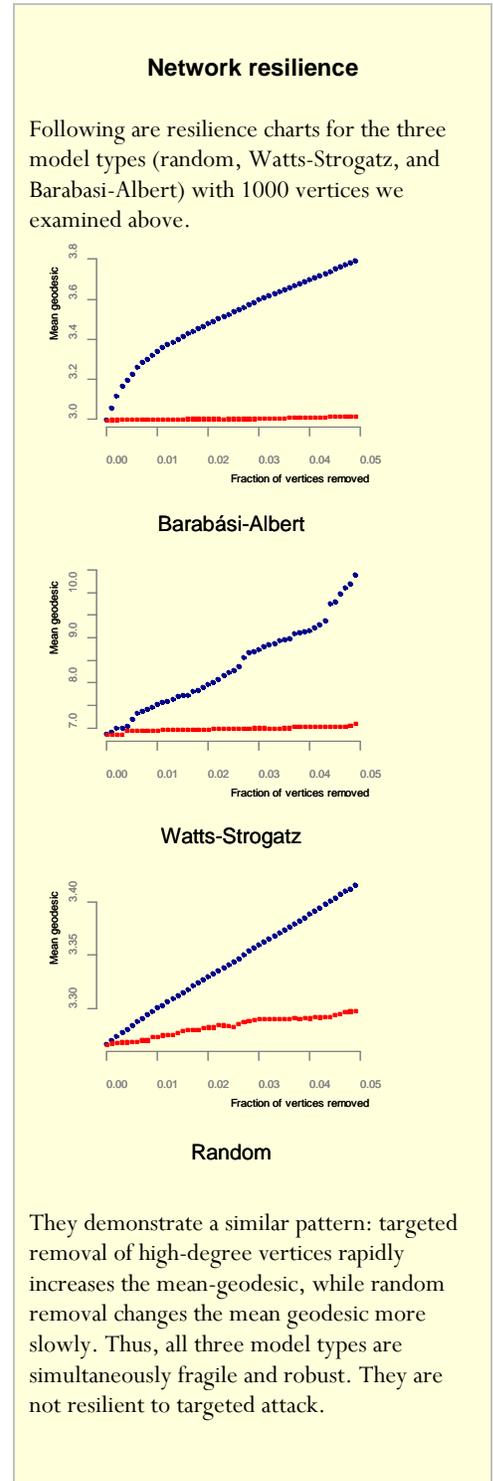
²⁷ Plotted with PlotNetworkResilience().

D. REAL-WORLD NETWORKS CONTINUED

As the chart demonstrates, because there is hardly any change in the mean geodesic as vertices are randomly removed, the airline network is resilient to random failure of its vertices (airports). It is robust to random failures.

But ordered removal is another matter. As the largest vertices are removed, the mean geodesic increases dramatically. With only 3 percent of its largest vertices removed, the mean geodesic more than triples, and effectively destroys the air travel network. This result brings a grave message: Because the airline network is not resilient to targeted airport removal, targeted terrorist attacks at key airports could incapacitate the network. It is fragile to targeted attack.

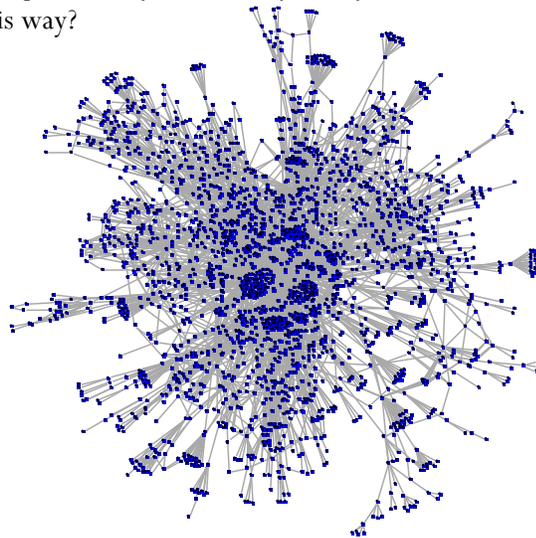
Unfortunately, as the sidebar shows, changing the network from a Barabasi-Albert model to a Watts-Strogatz model, or even to a random model, won't help with this problem: all three models are fragile to targeted attack.



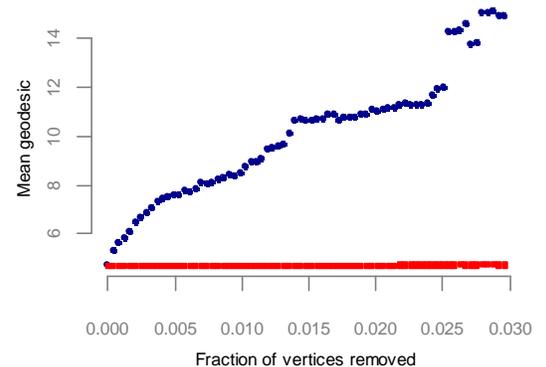
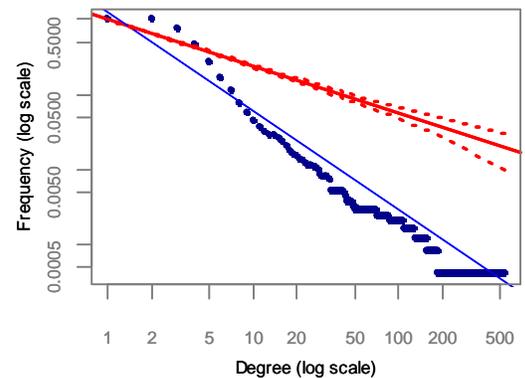
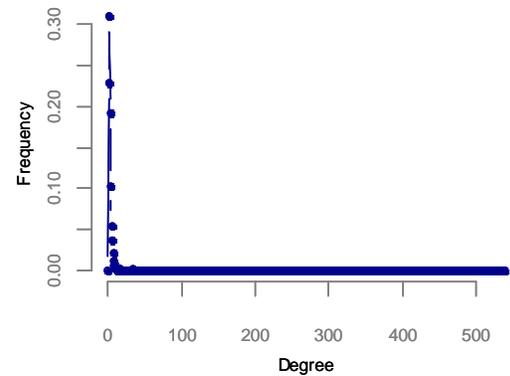
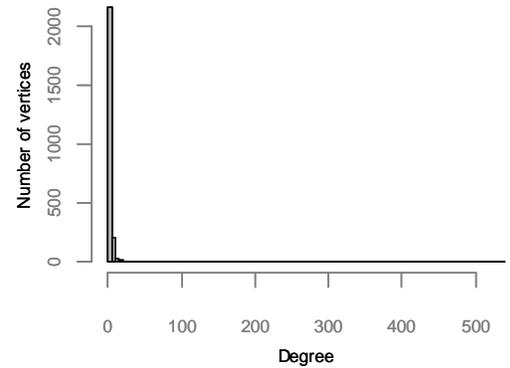
D. REAL-WORLD NETWORKS CONTINUED

2. Gene interactions

Now let's look at another real-world network, one from the natural world. A set of about 2,500 genes interact with one another to control all the metabolic interactions of our cell components. An overview of these genes and their interactions is shown in the charts and statistics on this page.²⁸ Given this information, what can you conclude about the gene network? Is it a small world? If so, what kind? Is it fragile or robust, or both? Is it power law distributed? What might the structure of this network say about our cells' resilience or susceptibility to disease? Is the network a complex system with emergent characteristics, or merely a complicated system? Why do you think our genes evolved in this way?



Number of vertices	2,433
Number of edges	4,642
Diameter	13
Mean degree	3.8
Maximum degree	539
Minimum degree	1
Mean geodesic	4.7 (random = 6.0)
Density	0.002
Transitivity	0.006 (random = .001)
Correlation coefficient	-0.125
Number of clusters	1



²⁸ Created with CreateGeneOntologyGraph() and PlotGraph().

E. PRACTICAL APPLICATIONS

In the preceding sections of this chapter, we focused primarily on understanding what networks are, their structures, their statistical characteristics, how they emerge, and the model types we can use to represent them. In the end, though, our goal is to understand how real-world networks behave. Ultimately, as actuaries, we want to better understand – and perhaps even predict – the behavior of social networks (such as insurance plan members, healthcare providers, and pension plan members), the behavior of organizational networks (such as insurance and reinsurance companies), and even the behavior of economic networks as a whole.

In the previous section, in our exploration of the global airline network, we saw how network theory can help us better understand the behavior and resilience of an important real-world system. Here we will look at further practical applications of networks that are of interest to actuaries. Keeping in mind that the study of complex system networks is barely ten years old, the results to date are impressive. The primary practical breakthroughs of interest to actuaries relate to the propagation of disease and information over networks, and to the assessment of network resilience.

1. Disease and information propagation

In 2000, Kuperman and Abramson used computerized network models to explore disease propagation, and made important discoveries.²⁹ First, they found that on a lattice-like social network, a disease never gets very far. For large populations, the fraction of people infected only reaches about one percent. However, on social networks like the Watts-Strogatz model, the fraction infected depends on the number of rewired vertices (those with long-distance relationships). When the percentage of such rewired vertices is ten to twenty percent, there is a sudden and dramatic transformation, a tipping point. At this point the disease can spread to as much as 30 percent of the population. These results were encouraging news for public health officials: they suggested that by keeping a disease below its tipping point (such as by mass immunization), a disease can be contained.

²⁹ See “Small world effect in an epidemiological model” in M.E.J. Newman, et al. (2006), pages 489-492.

E. PRACTICAL APPLICATIONS CONTINUED

Unfortunately, worse news came in 2001. The physicists Pastor-Satorras and Vespignani found that in a social network that follows a Barabasi-Albert model, with hubs and connectors, there is no tipping point.³⁰ Diseases on such a network will spread to a large percentage of the population, period. Such is the situation with AIDS. It turns out that the social sexual network of most societies is one of preferential attachment, with hubs and connectors having large numbers of edges (sexual contacts). On such a network, in the absence of a cure, AIDS cannot be stopped from spreading; it will never disappear. However, together with this sobering news comes a potential solution: as with the airline network, the way to solve the problem is to change the type of network. Random ‘vaccination’ (i.e., for AIDS, education and medical treatment) won’t work, but targeted ‘vaccination’ will. If we can identify and ‘vaccinate’ the hubs in the sexual network, AIDS could eventually die out.

Pastor-Satorras and Vespignani also studied computer viruses on the World Wide Web, and found the same result. Because the World Wide Web is a small world network with a power-law degree distribution, there is no epidemic threshold for computer viruses. No matter how short a time computers spend in an infected state, or how little they pass on the virus, the virus remains endemic. This result has been borne out by empirical research.

³⁰ See “Epidemic spreading in scale-free networks” in M.E.J. Newman, et al. (2006), pages 493-496

E. PRACTICAL APPLICATIONS CONTINUED

2. Network resilience

Speaking about the evolution of the world financial network, in a recent speech Andrew Haldane, Executive Director for Financial Stability of the Bank of England, said, “This evolution in the topology of the network meant that sharp discontinuities in the financial system were an accident waiting to happen. The present crisis is the materialisation of that accident.”³¹ (See the sidebar for details.)

Just as we saw with the airline network, networks such as the worldwide financial network, the Internet, and the World Wide web – all with high-degree hubs – are resilient to random system distress, but fail rapidly when the hubs fail.

Since the world financial system is integral to much of actuarial work, keeping track of the financial network’s topology may become an important concern for you. And what about the topology of the health insurance sector, the life insurance sector, the reinsurance sector? Do these have high-degree hubs? How fragile and susceptible to failure are they?

Similarly, in the domain of enterprise risk management, mapping the topology of communication and management relationships within an organization (perhaps by mapping email traffic) could help you understand its resilience. But as well, what you’ve learned about networks can be applied to operational actuarial work. For example, a health insurance actuary might employ networks to explore the potential for catastrophic disease spread among members, the potential efficacy of targeted vaccinations and other disease management efforts, and the structure of physician referral networks. What other applications can you think of for your organization?

Unfortunately, it appears that the specific application of networks to actuarial problems is currently unknown. In the literature search for this report, I used the search terms “network science” and “network theory” to find relevant resources on the major actuarial web sites. There was not a single relevant hit.³²

³¹ See Haldane (2009), and especially see the charts used in his speech.

³² See the section titled *Finding the essential resources* in this report.

An unholy trinity

In his speech, Andrew Haldane noted that over the past 20 years the world financial network has become:

- more dense (more vertices)
- more complex (more edges)
- more asymmetrical (more skewed degree distribution)
- a less separated “small world” (lower mean geodesic)

He further says, “So based on evidence from a sampled international network, the past twenty years have resulted in a financial system with high and rising degrees of interconnection, a long-tailed degree distribution and small-world properties. That is an unholy trinity. From a stability perspective, it translates into a robust-yet-fragile system, susceptible to a loss of confidence in the financial hubs and with rapid international transmission of disturbances.”

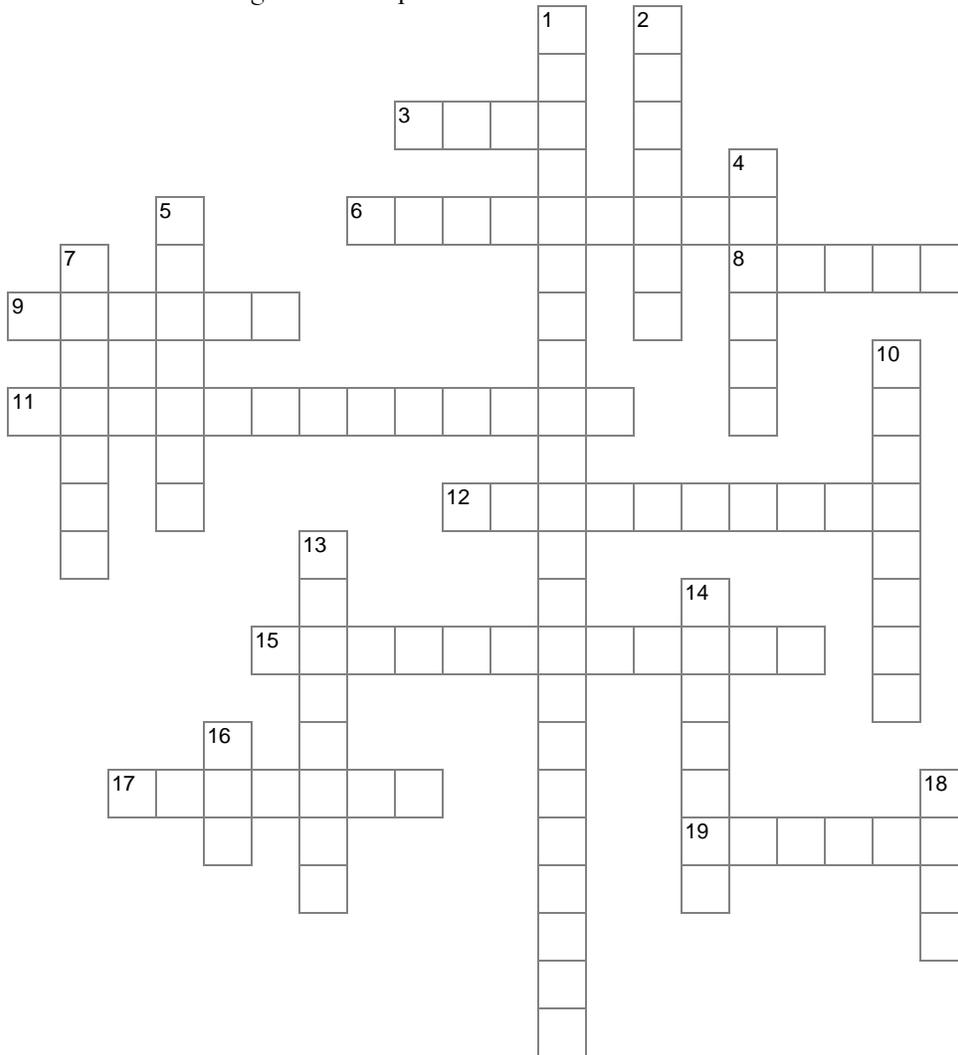
He suggests three policy prescriptions:

- Putting in place a system to map the global financial network and communicate to the public about its dynamics.
- Increasing regulation to better manage fragilities within the financial network, especially high-degree hubs.
- Restructuring the network to reduce the chances of financial collapse.

F. EXERCISES

1. Make sure you can reproduce each of the figures in this chapter, and verify each of the chapter’s numerical and formulaic results. For this exercise, review the chapter’s footnotes.

2. Solve the following crossword puzzle.



ACROSS

- 3 In a graph, the representation of a relationship
- 6 Characteristic of graphs that follow a power law distribution
- 8 A representation of a network
- 9 In a graph, the representation of an agent
- 11 Mixing by vertex degree (extra credit)
- 12 A graph with low mean geodesic and transitivity higher than a comparable random graph
- 15 When the friend of a friend is a friend
- 17 An island of vertices
- 19 Equivalent to power law

DOWN

- 1 What differentiates Barabasi-Albert graphs
- 2 He mailed packages and discovered the small world effect
- 4 The number of edges attached to a vertex
- 5 A real-world system of agents connected by relationships
- 7 Ratio of the number of edges to the possible number
- 10 Shortest path length
- 13 An edge with arrows
- 14 A directed graph
- 16 A high-degree vertex
- 18 An edge for a vertex connected to itself

F. EXERCISES CONTINUED

3. What is the maximum number of edges in an undirected graph with N vertices? Such a graph is called a ‘**complete N -graph**’.

4. In the Package ‘igraph’ reference manual, look up “grg.game” (geometric random graphs), and write an R function to generate a geometric random graph with 1000 nodes and a radius of 0.0575 (with the remaining arguments left at their defaults). Save the graph to a cvs file. Then, using PlotGraph(), plot it interactively, experimenting with different layouts. What does its plot look like to you in the ‘FR’ layout? Next, analyze the graph using AnalyzeGraph(). How does your graph compare to the random graph example with 1000 vertices shown in the report? Can you explain the differences?

5. The definition of a small world network involves two measures: one comparing a network’s transitivity to that of a random network, and one comparing a network’s mean geodesic to that of a random network. Can you develop one measure to determine whether a network is a small world? Using the table on page 10 of the paper “The structure and function of complex networks” by Newman³³, test your new measure. You should also test various random networks, to make sure that your measure differentiates a small world from a random world.

6. Think of a project within your firm for applying what you have learned in this chapter. Then, using your new tools and knowledge, carry out the project. If you like the results, please let me know what you did and what you discovered.

G. TO LEARN MORE

To learn more about networks, you may first enjoy watching the video “Connected: the power of six degrees”.³⁴ Then, read *Linked* by Albert-Laszlo Barabasi, one of this report’s *Top ten complexity science books*.³⁵

After that, you may find it interesting to read M.E.J. Newman (2008), M.E.J. Newman (2003), M.E.J. Newman, et al. (2006) – another of this report’s *Top ten Complexity Science books* – and Barrat, Barthelemy, & Vespignani (2008), perhaps in that order.

³³ M.E.J. Newman (2003)

³⁴ See Discovery Channel (2008), a five-part video on “www.youtube.com”.

³⁵ See the book’s annotation in the section of this report titled *Top ten Complexity Science books*.

H. REVIEW AND A LOOK AHEAD

This chapter introduced the first of four archetypal Complexity Science model types: networks. You learned what networks are and how they are represented by graphs. You also learned their basic terminology, four basic graph types (random, lattice, Watts-Strogatz, and Barabasi-Albert), network statistical measures such as the mean geodesic, transitivity, and correlation coefficient, different types of degree distribution, how the network topology of real-world networks affects our lives, and why real-world networks are often complex systems. You also learned how to use ‘R’ and ‘igraph’ to develop and analyze graphs.

I hope you enjoyed the introduction, and are sufficiently intrigued to consider applying what you have learned about networks in your work.

Next we will look at the archetypal Complexity Science model type that, upon an underlying network structure, adds agent behavior.

CHAPTER FIVE: CELLULAR AUTOMATA

I did what is in a sense one of the most elementary imaginable computer experiments: I took a sequence of simple programs and then systematically ran them to see how they behaved. And what I found – to my great surprise – was that despite the simplicity of their rules, the behavior of the programs was often far from simple. Indeed, even some of the very simplest programs that I looked at had behavior that was as complex as anything I had ever seen.

Stephen Wolfram¹

A. INTRODUCTION

Stephen Wolfram is a genius. He published his first physics paper at age 15, and two years later published another that caught the attention of Nobel prize winning physicist Murray Gell-Mann, who invited Wolfram to join his group at Caltech. Wolfram obtained a PhD in theoretical physics from Caltech at age 20 (the youngest person ever to receive such a degree), joined its faculty, and soon received one of the first MacArthur Foundation ‘genius’ grants. A couple of years later, he joined the prestigious Princeton Institute for Advanced Study.

About the time that Wolfram was finishing his PhD, he became interested in ‘**cellular automata**’ (see the sidebar). Since then, he has spent most of his professional life studying them. At first blush, cellular automata seems a queer subject for one of the world’s most brilliant physicists to bother about at all, much less for most of his life. (As you will see, the subject is not even remotely like the typical physics topics.) Yet, speaking about his work with one of the most simple cellular automata – what he calls the ‘rule 30’ cellular automaton – he says this was “probably the single most surprising scientific discovery I have ever made”.³ Indeed, Wolfram now thinks that cellular automata hold the key to understanding our universe. His 2002 book *A new kind of science* – one this report’s *Top ten Complexity Science books* – is a 1200-page paean to cellular automata that details how they enlighten most areas of human inquiry, even philosophy.

This chapter introduces cellular automata, our second model archetype, which includes agents, agent relationships, and agent behavior. You will learn how simple agent behavior can lead to complex system-wide results. You will also learn what cellular automata are, the basics of behavior rules, four common model types, and how they are applied. Along the way, perhaps you too will become entranced by them, or at least see how they can enlighten actuarial work.

Cellular automata

In the 1950s, John von Neumann, another genius, began wondering if a self-operating robot-like machine (an ‘automaton’) could manufacture itself. At that time, the idea of robots – especially self-replicating robots – was entirely new.

Von Neumann suspected that self-replicating robots were possible, but with the technology then available, he thought building one would be impractical. Mathematician Stanislaw Ulam suggested an alternative: that Von Neumann experiment with such machines in an abstract universe, a two-dimensional grid of cells with various states representing machine parts.

Von Neumann liked Ulam’s idea, and thus ‘cellular automata’ were born. He created a cellular automaton that could not only reproduce itself (thus proving the possibility of self-replicating robots), but could also reproduce any pattern placed on a two-dimensional cellular grid. He also showed that his self-reproducing automaton was a universal Turing machine.²

But it was not until Stephen Wolfram began his studies, some thirty years later, that cellular automata became recognized as a vitally important subject. It is now an established scientific discipline with applications spanning nearly every field of science. Referencing Wolfram’s early work are now more than 10,000 published papers.

¹ Wolfram (2002), page 2. This book is one of this report’s *Top ten Complexity Science books*.

² Von Neumann, J. *Theory of self-reproducing automata* (edited and compiled by A. W. Burks). Urbana: University of Illinois Press, 1966.

³ Wolfram (2002), page 27.

B. BASICS

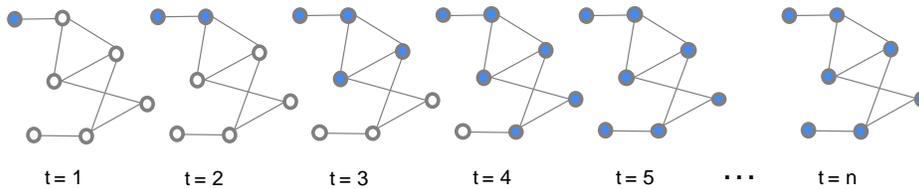
Let's first define a 'cellular automaton' (abbreviated 'CA'; plural 'cellular automata'). Consider a graph with:

- vertices that can each assume two or more 'states' (think of various colors representing the states), with the state of each vertex depending on its previous states and on the states of other vertices in a local neighborhood, and
- a 'behavior rule' governing how each vertex's state is updated.

A CA is such a graph, incorporating the allowable vertex states, and the behavior rules governing the CA's evolution.⁴

For example, following is the evolution of a CA over n time steps. For each of this CA's vertices, there are two possible states, represented by colors blue and white. Its global behavioral rule is:

- If a vertex was blue at the previous time step, it remains blue.
- If at the previous time step, a next neighbor (a vertex within a radius of one edge) of a white vertex was blue, the vertex turns blue.



For the CA above, if we represent the white state by 0 and the blue state by 1, the behavioral rule can be written:

$$\text{if } \sum_{i \in N} v_{t-1}^i \geq 1 \text{ then } v_t^x = 1; \text{ else } v_t^x = 0$$

where N is the set of all vertices within a radius of one edge from vertex x, and v_t^i is the state of vertex i at time t.

The example above is a 'first-order CA', because its behavior rule for time step t only depends on states at the previous time step t-1. The behavior rule of a 'second-order CA' depends on states at time steps t-1 and t-2. Also, the example's behavior rule has a radius of 1. Behavior rules with higher radii are also possible, but generally the behavior rule radius is much less than the graph's diameter.

⁴ Although CA are typically defined on regular lattice graphs (such as 1-, 2-, and higher-dimensional grids) here we are introducing a more general definition that encompasses newer model types.

Complexity science – an introduction (and invitation) for actuaries

Chapter five: Cellular automata continued

B. BASICS CONTINUED

In the example above, the behavior rule is deterministic; chance doesn't play a role. However, the second part of the rule could have been changed to "If at the previous time step, a next neighbor of a white vertex was blue, the vertex turns blue with a probability of 0.50. Otherwise, it remains white". Such a behavior rule is stochastic; it depends on chance.

In the example above, the behavior rule is applied to each vertex at each time step. Such updating is called '**synchronous updating**'. If, however, at odd time steps only half of the vertices were updated, and at even time steps the other half were updated, such an updating scheme would be an example of '**asynchronous updating**'.

As you learned in Chapter two, behavior rule algorithms can range from simple if-then statements to complicated optimization algorithms.

The following table summarizes the characteristics of the four CA model types that we will explore in this chapter.

Model type	Graph			Behavior rule			
	type	boundary	vertex states	type	radius	order	updating
1. Simple 1D⁵	1D lattice	periodic	2	deterministic	1	1	synchronous
2. Simple 2D	2D lattice	periodic	2	deterministic	1	1	synchronous
3. Sandpile	2D lattice	cliff	5	stochastic	1	1	synchronous
4. Forest fire	2D lattice	periodic	4	stochastic	1	1	synchronous

A CA '**boundary**' is applicable only to CA on a lattice graph. It refers to how cells at the edge of the lattice are related to other cells. Later examples will make this concept clear.

That's it. Those are the CA basics. That is, in essence, what Wolfram has spent his life studying. However, if you think that so few basics can hardly lead to anything interesting, prepare yourself for quite a surprise.

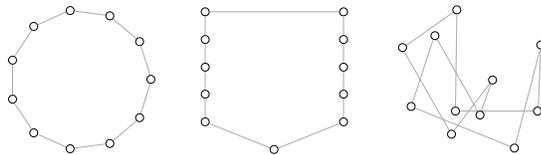
⁵ '1D' is an abbreviation for 'one-dimensional'.

C. CA MODELS

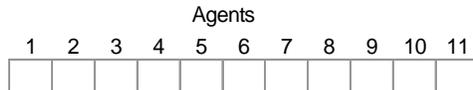
There are four common CA model types: ‘Simple one-dimensional CA’, ‘Simple two-dimensional CA’, the ‘Sandpile model’, and the ‘Forest fire model’. This section describes each of these model types.

1. Simple 1D-CA

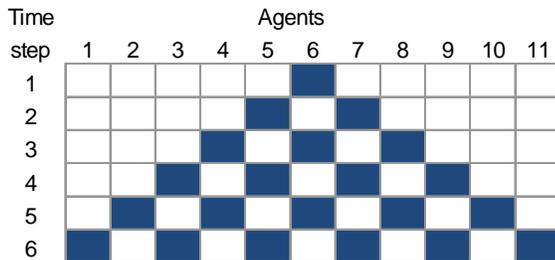
The simplest CA model type, called a ‘**simple one-dimensional CA**’ (1D-CA), is a graph whose vertices all have exactly two nearest neighbors (ie, neighbors within a ‘radius’ of 1). As you saw in the last chapter about networks, such a graph can be drawn in many layouts.



But the layout most commonly used for 1D-CA is a linear grid where each cell is a vertex (in the CA field more commonly called a ‘cell’ or an ‘agent’ – we will call it an agent), and where the first and last cells are understood to connect (a so-called ‘**periodic boundary**’).



To show the evolution of a simple 1D-CA over time, a two-dimensional grid is used, where a column of the grid represents the series of states for an agent, and a row represents the graph’s ‘configuration’ at a particular time step. The grid arrangement makes it easy to visualize the pattern of the graph’s configurations as it evolves.



Thus, the grid above shows the evolution of 11 agents over 6 time steps, where each agent can have two states (represented by the white and dark colors).

C. CA MODELS CONTINUED

1. Simple 1D-CA continued

The first class of simple 1D-CA that we will study has the following characteristics:

- the graph is a one-dimensional lattice with a periodic boundary, including vertices (agents) with two states
- behavior rules are deterministic, first order, with a radius of 1, and with synchronous updating.

For these CA, the behavior rule to determine the state of an agent at time t is based on the state of three agents at time $t-1$: the agent itself and its two neighbors. There are eight, 2^3 , possible arrangements of these three agents:



(Note that these states from left to right are the negative images of the states from right to left.)

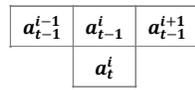
A clear way to show the behavior rule is to show the state of an agent underneath the eight possible states of its neighbors at the previous time step:



Thus, the behavior rule above provides that an agent changes its state for four of the previous state configurations; otherwise, the agent doesn't change its state. This is the behavior rule for the example CA with 6 time steps shown on the previous page. Can you verify that the behavior rule produces the example CA?

A second way to represent this behavior rule is a 'transition table', where '1' represents the dark color, '0' represents white, and a_t^i represents the state of agent i at time t :

a_{t-1}^{i-1}	a_{t-1}^i	a_{t-1}^{i+1}	a_t^i
1	1	1	1
1	1	0	1
1	0	1	1
1	0	0	1
0	1	1	1
0	1	0	0
0	0	1	1
0	0	0	0



C. CA MODELS CONTINUED

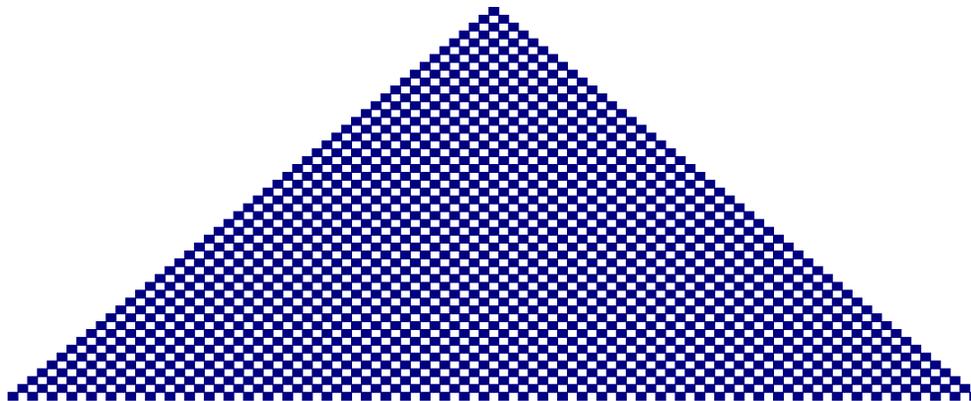
1. Simple 1D-CA continued

And a third way to represent the behavior rule is a ‘**rulestring**’, a string of 1’s and 0’s corresponding to the last column of the transition table. The rulestring for the behavior rule given above is thus ‘11111010’ (the last column of the transition table, from top to bottom), which in base 2 arithmetic equals the decimal number 250:

$$250 = 1*2^7 + 1*2^6 + 1*2^5 + 1*2^4 + 1*2^3 + 0*2^2 + 1*2^1 + 0*2^0$$

We can therefore refer to the behavior rule for our example CA as the ‘rule of 250’. Even more simply, the CA as a whole can be referred to as the ‘1D-CA-250’.

Below are 50 time steps of 1D-CA-250 with 100 agents, and with only the central agent (agent number 51) initially in the ‘on’ state (ie, with a state of 1). To make the CA’s pattern easier to see, grid lines around the lattice cells are turned off.⁶

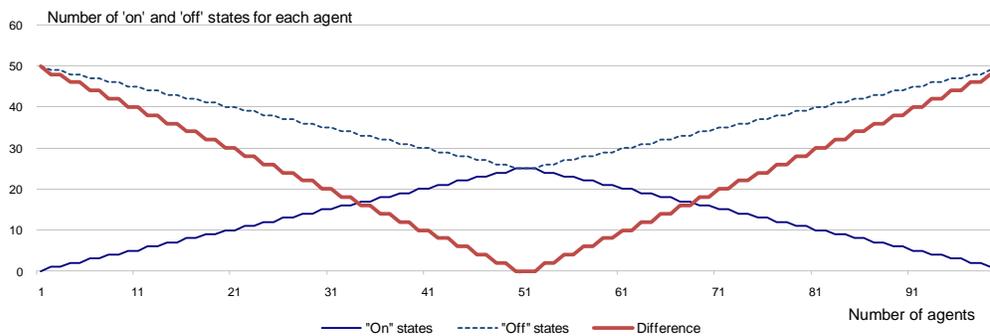
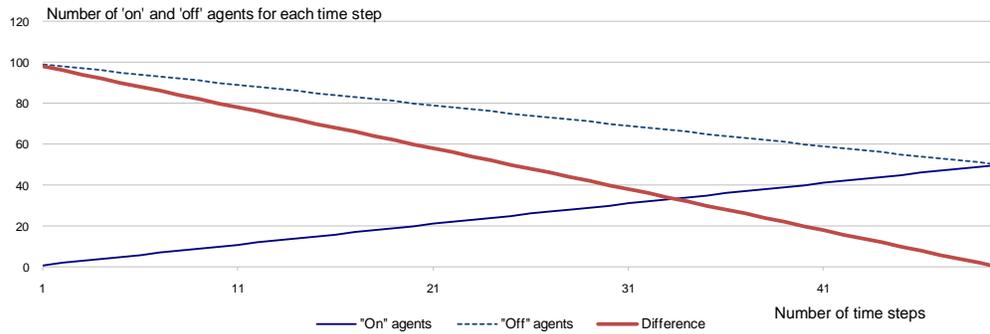


⁶ Created with the Excel one-dimensional CA model, with the following parameters: behavior rule: 250; number of agents: 100; number of time steps: 50; agents initially ‘on’: one agent – center.

C. CA MODELS CONTINUED

1. Simple 1D-CA continued

Let's continue to explore the 1D-CA-250. Following are two charts analyzing its behavior:⁷



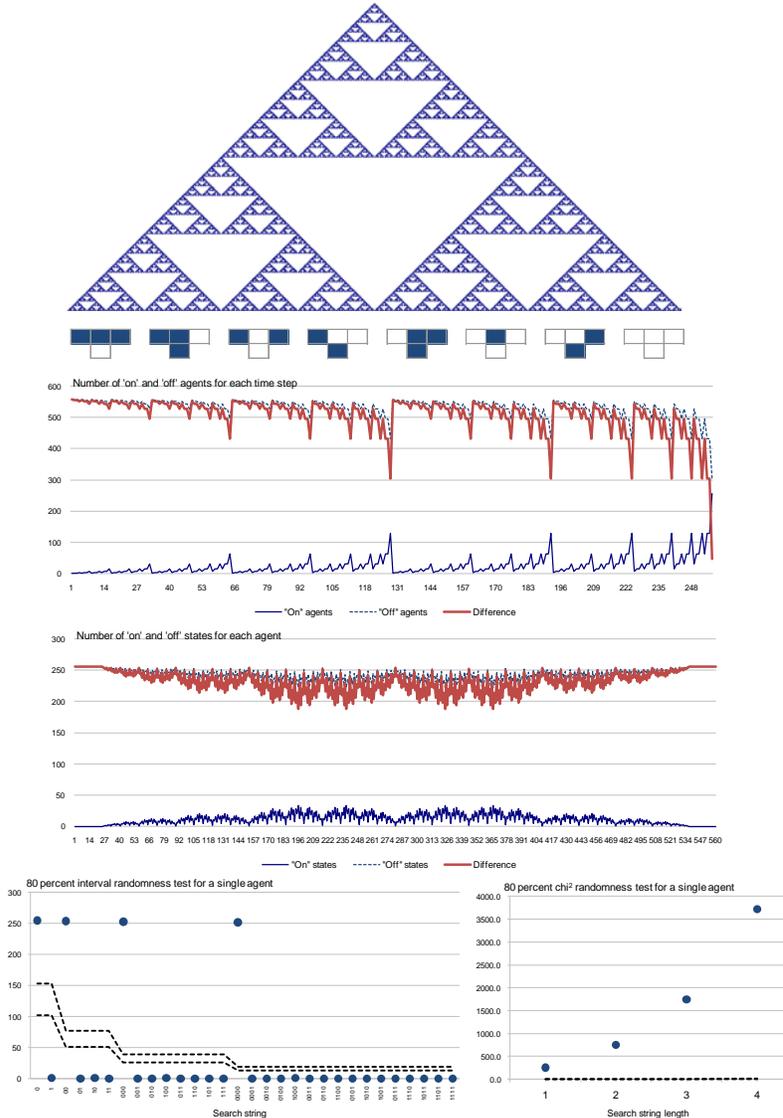
The top chart shows the number of agents in the 'on' and the 'off' states at each time step (there are 50 time steps), together with the difference in these numbers. The second chart shows the total numbers of 'on' and 'off' states, and their difference, for each agent over the 50 time steps (there are 100 agents). Can you verify these results?

⁷ These charts are found on the Analysis sheet of the Excel workbook used to create the CA.

C. CA MODELS CONTINUED

1. Simple 1D-CA continued

Following is 1D-CA-90, for 560 agents and 256 time steps.¹⁰

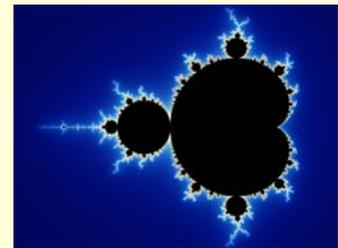


Even though its behavior rule is again quite simple, this CA is more interesting. Perhaps most striking is its nested ‘**fractal**’ pattern, made up of triangular areas each of which is a smaller copy of the whole pattern. (It is one case of the so-called Sierpinski triangle.) But, as the charts show, even though the number of ‘on’ states at each step is rather volatile, its pattern is again highly regular, not at all random.

Fractals everywhere

In the 1970s Benoit Mandelbrot, French-American mathematician, pondered the clouds, snowflakes, mountain ranges, trees, coastlines and similar forms that make up our everyday experience, and wondered if there is a way to describe them mathematically. Clearly, they are not made of regular circles, triangles, or squares. But what are they?

To his great surprise he found that most complex forms of nature could be described by very simple recursive mathematical formulas that he called fractals (because they produce shapes that are ‘fractured’, like in the real world). An example is the formula for the famous Mandelbrot set: $Z = z^2 + c$



He also made another important discovery: that such fractal forms, like many forms in nature, are ‘self-similar’, that they look generally the same at different scales of magnification, but that they exhibit new detail at all scales. No matter how closely you magnify a fractal, you see details that you did not see before. And, importantly, such details are not predictable. Thus, in the completely deterministic world of fractals, prediction is not possible.

For a fascinating overview of fractals, see the video narrated by Arthur C. Clarke, with Benoit Mandelbrot, Stephen Hawking and others: *Fractals – the colors of infinity*.⁹

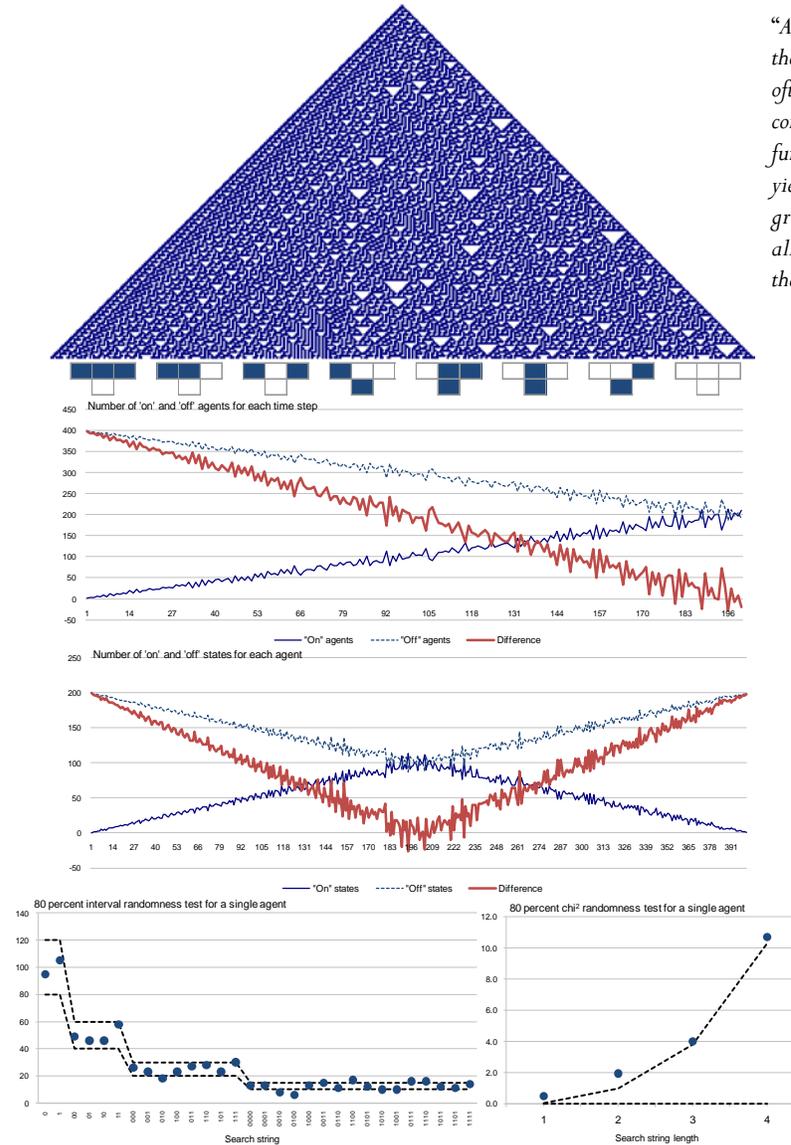
⁹ Clarke (2008). You may also enjoy the beautiful book Mandelbrot (1983).

¹⁰ Created with the Excel one-dimensional CA model, with the following parameters: behavior rule: 90; number of agents: 560; number of time steps: 256; agents initially ‘on’: one agent – center; individual agent to analyze: 281.

C. CA MODELS CONTINUED

1. Simple 1D-CA continued

Now let's look at 1D-CA-30 for 400 agents and 200 time steps.¹¹



“And what I found – to my great surprise – was that despite the simplicity of their rules, the behavior of the programs was often far from simple. . . . It took me more than a decade to come to terms with this result, and to realize just how fundamental and far-reaching its consequences are. . . . it yields a resolution to what has long been considered the single greatest mystery of the natural world: what secret it is that allows nature seemingly so effortlessly to produce so much that appears to us so complex.”

Stephen Wolfram
A new kind of science
 page 2

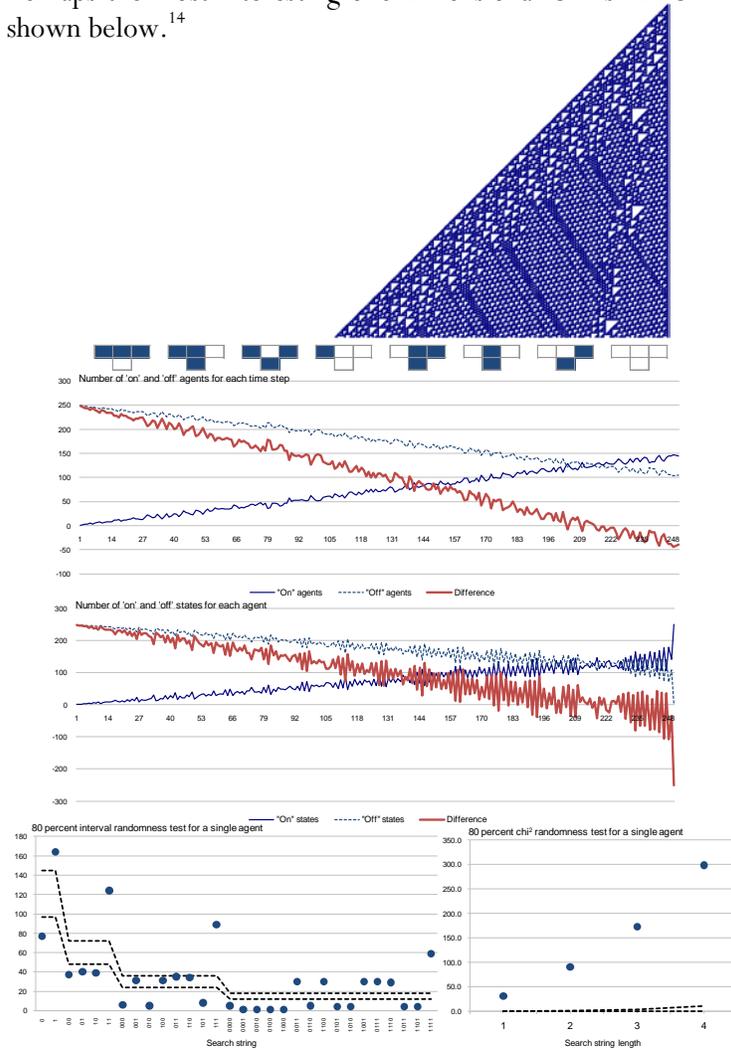
This is the CA that, when he first saw it, was “the single most surprising scientific discovery” that Wolfram had ever made. Given the regular patterns of the preceding CAs, and our intuitive expectations about the effects of simple rules, this CA is indeed startling. Its pattern is anything but regular. In fact, it is perfectly random, so much so that Wolfram patented its use as part of a pseudo-random number generator.

¹¹ Created with the Excel one-dimensional CA model, with the following parameters: behavior rule: 30; number of agents: 400; number of time steps: 200; agents initially ‘on’: one agent – center; individual agent to analyze: 201.

C. CA MODELS CONTINUED

1. Simple 1D-CA continued

Perhaps the most interesting one-dimensional CA is 1D-CA-110, shown below.¹⁴



In 1998, Matthew Cook, a research scientist working with Wolfram, presented a proof that this CA is a ‘**universal computer**’. That is, it can reproduce *any* computation your computer can do, and more. It is perhaps the simplest universal computer. This characteristic of rule 110 is the basis for Wolfram’s Principle of Computational Equivalence, perhaps the greatest achievement of *A new kind of science* (see sidebar).

A new law of nature

Wolfram says that the Principle of Computational Equivalence is a new law of nature: “Among principles in science the Principle of Computational Equivalence is almost unprecedentedly broad – for it applies to essentially any process of any kind, either natural or artificial. And its implications are both broad and deep, addressing a host of longstanding issues not only in science, but also in mathematics, philosophy and elsewhere.”¹²

The Principle states that:

1. **Everything is a computation.** Nature is a computer following simple rules, and the outcomes we see – complex as they may seem – are merely the result of nature’s computational process.
2. **Nature’s computations are equivalent.** Almost all processes that are not obviously simple are computations of equivalent sophistication.
3. **Universal computers are equivalent to nature.** Universal computers, such as rule 110, are equivalent to the most complex computations in nature.

“And what this suggests is that a fundamental unity exists across a vast range of processes in nature and elsewhere: despite all their detailed differences every process can be viewed as corresponding to a computation that is ultimately equivalent in its sophistication.”¹³

Although this new law is abstract in the extreme, it – if true (there is no consensus yet about the law’s truth, or even its meaning) – has serious practical consequences for actuaries (turn the page to find out).

¹² Wolfram (2002), page 715.

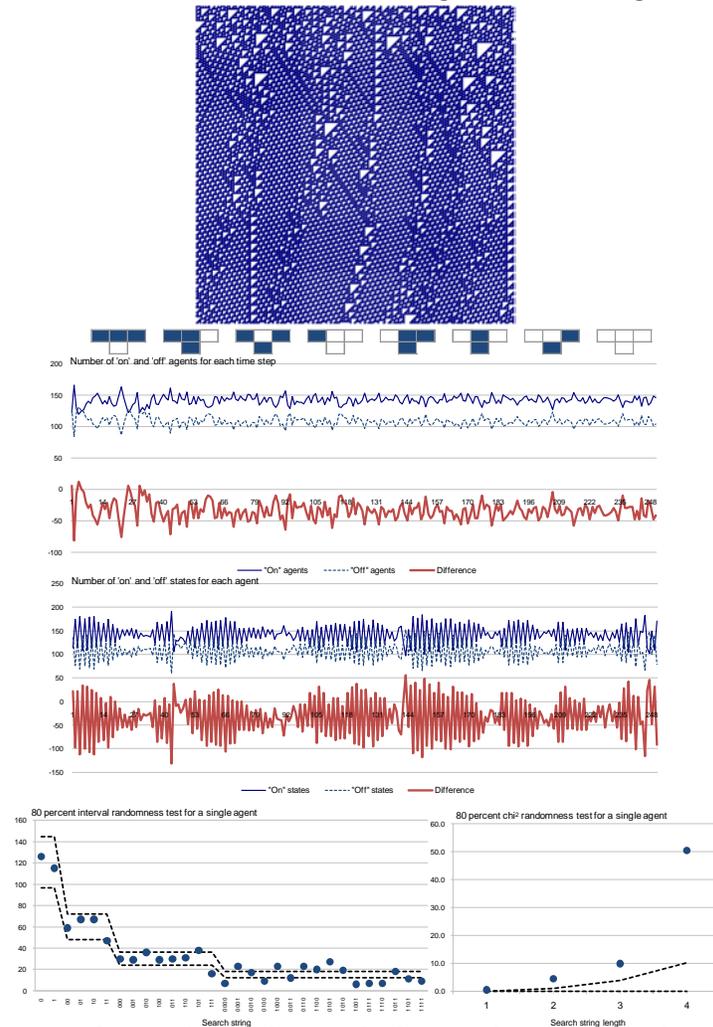
¹³ Wolfram (2002), page 719.

¹⁴ Created with the Excel one-dimensional CA model, with the following parameters: behavior rule: 110; number of agents: 250; number of time steps: 250; agents initially ‘on’: one agent – right; individual agent to analyze: 240 for time steps 10-250.

C. CA MODELS CONTINUED

1. Simple 1D-CA continued

Let's take another look at 1D-CA-110, this time with the initial 'on' agents at the first time step being uniformly randomly distributed, rather than simply starting with one 'on' agent.¹⁵



Now the CA's evolution looks much more random and unpredictable. Indeed, if you were given the state history of any one agent and did not know the underlying behavior rule, you could not predict its future (ie, you could not find an algorithmic shortcut to the process of rule 110 itself). Under the Principle of Computational Equivalence (if true) this implies that our attempts to predict complex social phenomena may be futile. The only way to find out the future is to watch the process unfold.

¹⁵ Created with the Excel one-dimensional CA model, with the following parameters: behavior rule: 110; number of agents: 250; number of time steps: 250; agents initially 'on': many agents – randomly distributed; individual agent to analyze: 240 for time steps 10-250.

C. CA MODELS CONTINUED

1. Simple 1D-CA continued

Wofram suggests that the behavior of all CA (not just simple one-dimensional CA) can be classified into four classes in increasing order of complexity:

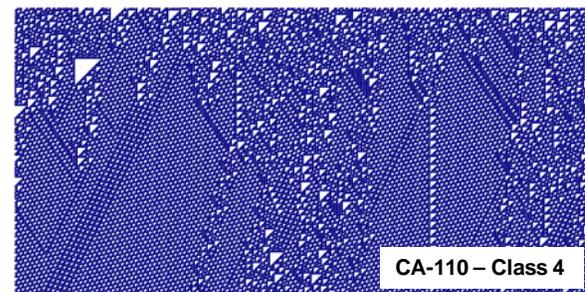
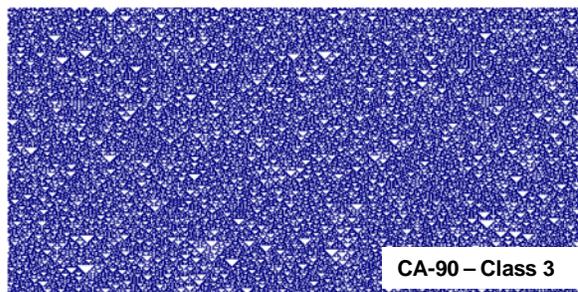
Class 1: The behavior of these CA is very simple, with almost all initial conditions evolving to a uniform final state.

Class 2: The behavior of these CA leads to simple structures that, although not uniform, are highly regular, either remaining the same forever, or repeating every few steps.

Class 3: These CA have more complicated behavior. They lead to configurations that are largely random, but with some islands of regularity (such as triangles).

Class 4: These CA combine order and randomness: their behavior produces some structures that are simple but that move around and interact with one another. They are poised between perfect organization (Class 2) and perfect randomness (Class 3), an important region sometimes called ‘**the edge of chaos**’ or ‘**critical**’, where many complex systems reside, and where universal computation, power laws, and catastrophic surprises are found.

Below are one-dimensional CA in each of these classes, with randomly distributed initial ‘on’ agents:¹⁶

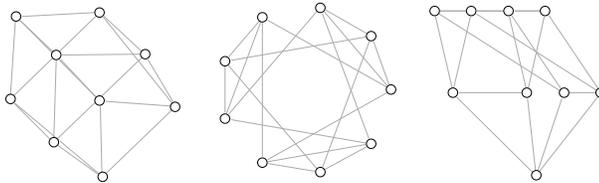


¹⁶ Created with the Excel one-dimensional CA model, with the following parameters: 500 agents, 250 steps, “many agents-randomly distributed”, random number seed = 2, and rules 250, 108, 90, and 110.

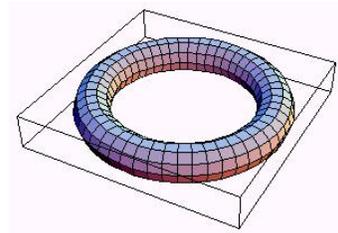
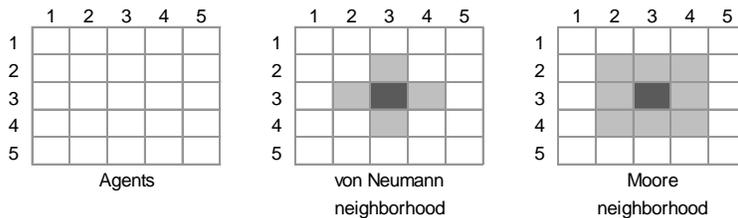
C. CA MODELS CONTINUED

2. Simple 2D-CA

A ‘simple two-dimensional CA’ (2D-CA) is a graph whose vertices all have either exactly four (von Neumann neighborhood) or exactly eight (Moore neighborhood) nearest neighbors. Such a graph can be drawn in many layouts. For example, a von Neumann 2D-CA with 9 vertices might be drawn as:



But the layout commonly used for 2D-CA is a two-dimensional grid where each cell is a vertex (or, an ‘agent’), and where opposite edges (east/west and north/south) are understood to connect to form a torus (a so-called ‘periodic boundary’).



The grid for a 2D-CA shows the CA’s configuration at one point in time. To show the evolution of a 2D-CA over time would require a three-dimensional figure.

The first class of 2D-CA that we will study has the following characteristics:

- agents have a von Neumann neighborhood and two states
- behavior rules are deterministic, of first order, with a radius of 1, and with synchronous updating

For these 2D-CA, the behavior rule to determine the state of an agent at time t is based on the state of five agents at time $t-1$: the agent itself and its four neighbors. There are 32, 2^5 , possible arrangements of these five agents at time $t-1$. These are commonly reduced to 10 combinations, as in the following transition table.

C. CA MODELS CONTINUED

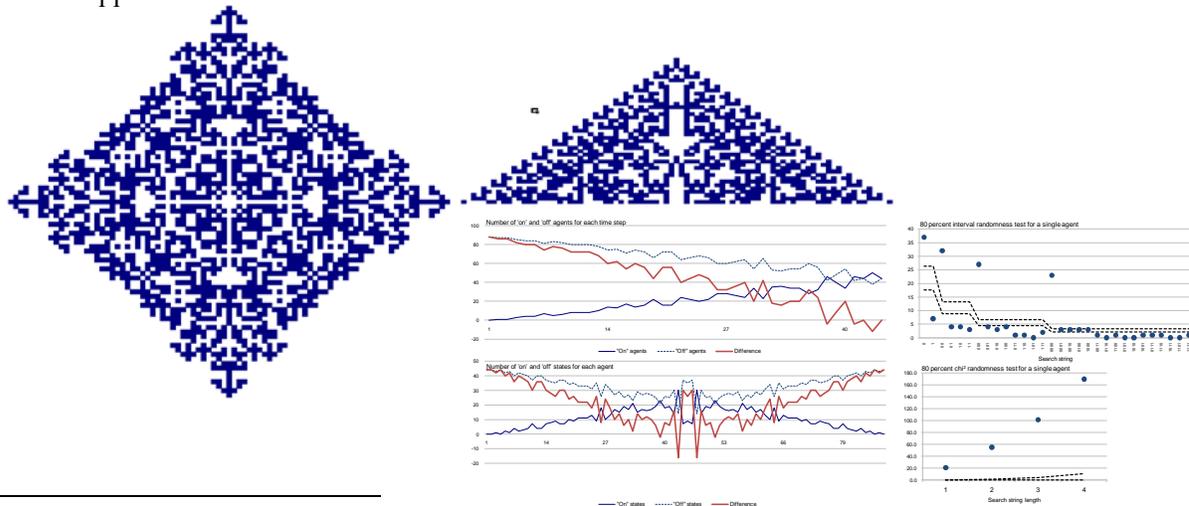
2. Simple 2D-CA continued

a_{t-1}^i	Number of 'on' agents in the neighborhood	a_t^i
1	4	0
0	4	1
1	3	1
0	3	1
1	2	0
0	2	0
1	1	1
0	1	1
1	0	1
0	0	0

The rulestring for the CA shown in the transition table is 462:

$$462 = 0*2^9 + 1*2^8 + 1*2^7 + 1*2^6 + 0*2^5 + 0*2^4 + 1*2^3 + 1*2^2 + 1*2^1 + 0*2^0$$

Therefore, we can refer to this CA as 2D CA-462. Below on the left is the 44th time step of 2D-CA-462 with 88x88 agents, and with only the central agent initially in the 'on' state. (To make the CA's pattern easier to see, the grid's lines around cells are turned off.)¹⁷ To the right is the time series of a cross section through the CA's middle row, together with analytic charts of the cross-section and of the center agent of the cross section. Does the cross section appear to be random?



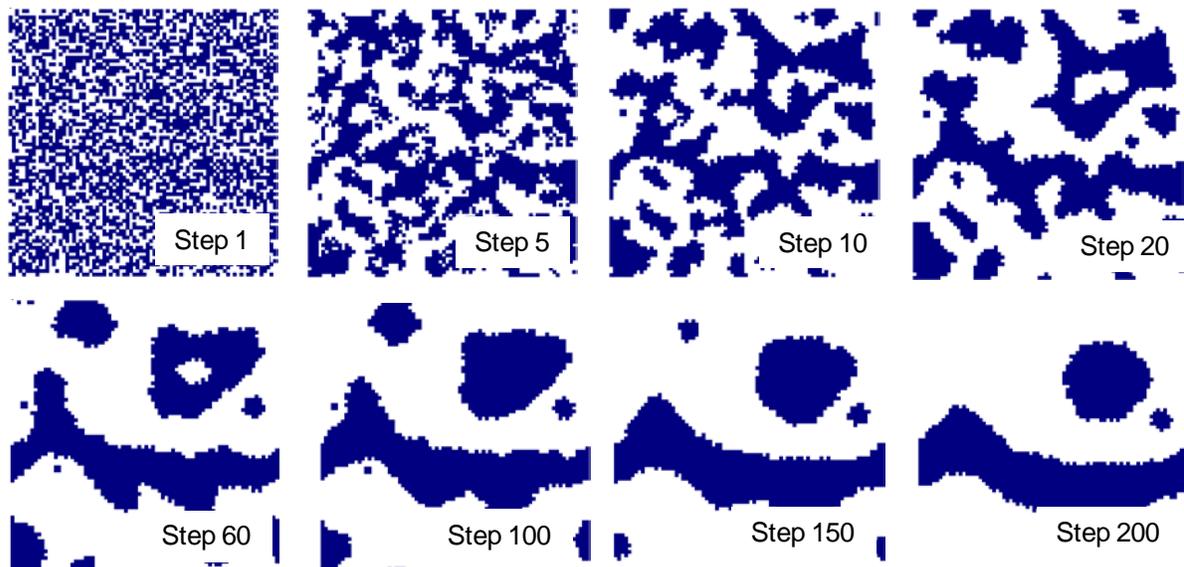
¹⁷ Created with the Excel 2D CA model, with the following parameters: behavior rule: 462; number of rows: 88; number of columns: 88; number of time steps: 44; agents initially 'on': one agent – centered; row to analyze: 44; column to analyze: 44. This is almost the same as the 2D-CA-462 on page 174 of Wolfram (2002), except that this 2D-CA would have to run one more time step to be exactly the same (because Wolfram counts his first time step as time 0, whereas I count it as time 1).

C. CA MODELS CONTINUED

2. Simple 2D-CA continued

The 2D-CA above is an example of a CA following a behavior rule with a von Neumann neighborhood. Following is an example of a CA with a Moore neighborhood.

It is 2D-CA-1,119,275,885,101,440, shown at intervals over 200 time steps.¹⁸



This sequence of CA configurations demonstrates that, even though CA evolve in many discrete steps, their overall pattern can evolve slowly in a manner that appears continuous. And it again demonstrates the impossibility of predicting how a CA will evolve.

It also shows how CA can self-organize out of a chaotic pattern into a structure that has almost life-like features and movements. We will next explore a 2D-CA that for more than 30 years has entranced people with its life-like patterns – The Game of Life.

¹⁸ Created with the Excel 2D CA model, with the following parameters: behavior rule: 1,119,275,885,101,440; number of rows: 80; number of columns: 80; number of time steps: 200; agents initially ‘on’: “Many ‘on’ agents – randomly distributed”; initial distribution percent: 50; random number seed: 2. This is the 2D-CA shown on p. 336 of Wolfram (2002). Wolfram’s rule number for this 2D-CA is 976, because he uses a highly aggregated rulestring algorithm for 2D-CA with Moore neighborhoods.

C. CA MODELS CONTINUED

2. Simple 2D-CA continued

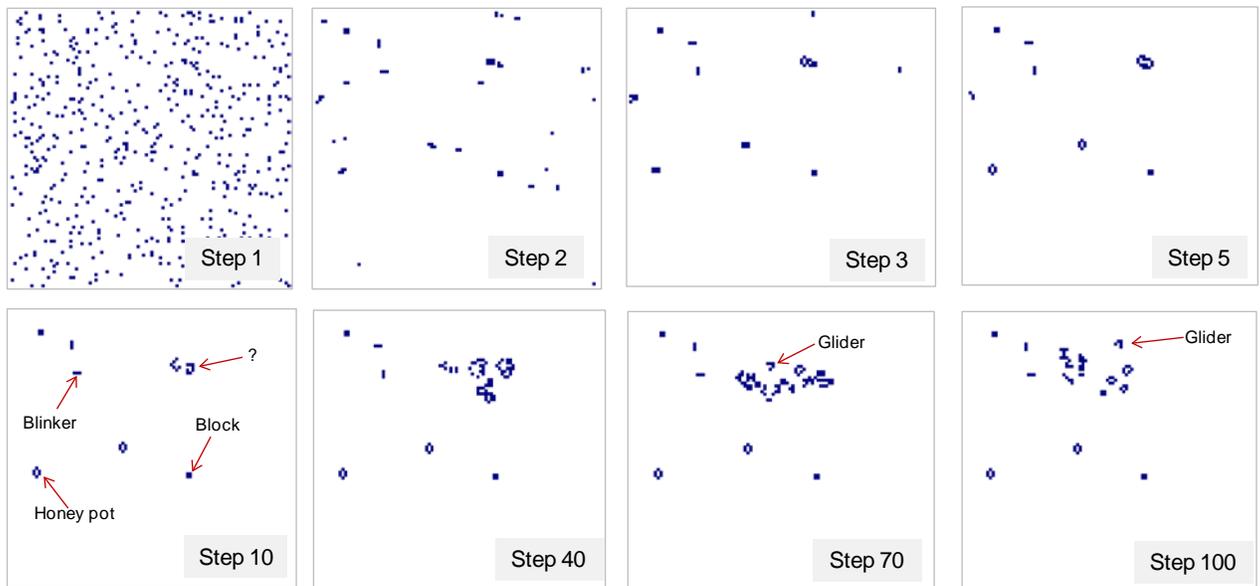
The Game of Life is actually not a game; rather, it is a 2D-CA with a simple behavior rule (see sidebar) that can evolve into a virtual world as rich as our own. The following sequence of CA configurations shows how the Game of Life (2D-CA-3,235,963,104) evolves from a 5 percent starting density of randomly distributed ‘on’ agents.¹⁹ One of the first surprises is that some Life objects move. Even though some don’t move (the so-called ‘still-lifes’ such as the ‘blocks’ and ‘beehives’ in the sequence below), and some simply alternate their states (such as the ‘blinkers’ below), some move or even explode, such as the exploding constellation in the upper right of the configurations below that produces a ‘glider’ taking off in the NE direction. Conway showed that it is possible to construct a universal computer simply out of ‘gliders’ and ‘blocks’.

The rule of Life

The Game of Life’s behavior rule is simple:

- An ‘off’ or ‘dead’ agent (state = 0) becomes ‘on’ or ‘alive’ if exactly three of its eight neighbors (in a Moore neighborhood) were alive during the last time step.
- A live agent (state = 1) stays alive if either two or three of its neighbors were alive during the last time step. Otherwise it dies.

In anthropomorphic terms, agents are born from *ménages à trois*. If they have only one companion, they die of loneliness, and if they have more than three they die from overcrowding.



Inherent in Life is mystery. Predictable as it is cell-to-cell and step-to-step, its overall pattern of evolution defies both our intuition and our forecasting mathematics. It is another example of a complex phenomenon arising from a simple rule, for which the only way to predict its evolution is to simply let it run.²⁰

¹⁹ Created with the Excel 2D CA model, with the following parameters: behavior rule: 3,235,963,104; number of rows: 100; number of columns: 100; number of time steps: 100; agents initially ‘on’: “Many ‘on’ agents – randomly distributed”; initial distribution percent: 5; random number seed: 2.

²⁰ For a riveting exploration of The Game of Life and its implications, you may enjoy reading Poundstone (1985). To see hundreds of varieties of ‘life forms’ in the Game of Life, you may enjoy visiting www.EricWeisstein.com/encyclopedias/life. You may also enjoy Conway’s presentation at Conway (2007a).

C. CA MODELS CONTINUED

3. Sandpile

In the late 1980s, the Danish physicist Per Bak became interested in systems that are in what physicists call a state of ‘**criticality**’, a state where a system is way out of balance, but not yet chaotic. It is in this state when systems experience ‘**phase transition**’ (such as ice turning to water, or non-magnetic material turning magnetic) and where their attributes exhibit fractal or power-law scale-invariance. Per Bak wondered how nature attains such criticality.

At that time, physicists thought that, in order to reach criticality, systems had to be manipulated, finely tuned. For example, the temperature of water had to be exactly 100° C or above for it to transition to steam, and the temperature of metals had to be below the ‘Curie point’ to transition to a magnetized state.

Contrary to this prevailing wisdom, Per Bak discovered something quite different. He found that systems will self-organize to a critical state, without outside manipulation, and he called his discovery ‘**self-organized criticality**’. During the ten years after Bak published the paper introducing this result, it was the most cited paper in physics. But this concept is also important outside of physics: it has profound implications for the social systems in which actuaries work.

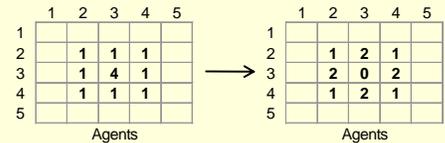
Self-organized criticality is demonstrated by Bak’s Sandpile model. The Sandpile model is a 2D-CA on a lattice with a ‘cliff’ boundary; that is, the lattice is finite, like a table top: agents at the edges are connected to only three or two (for agents at the corners) other agents in a von Neumann neighborhood, not to four. Each agent on the lattice has one of five states: 0, 1, 2, 3, or 4.

The behavior rule is first order with radius one, but includes an element of chance (see the sidebar).

The Sandpile model behavior rule

When the Sandpile model starts, at each time step a randomly-selected agent’s state is increased by one. For example, if the agent’s previous state was 2, it is increased to 3. This process continues until one agent’s state is 4. This is like dropping a grain of sand on a table-top grid.

At the next time step, any agent with a state of 4 (ie, with a stack of four grains of sand) ‘topples’; it resets itself to a state of 0, and adds one to each of its von Neumann neighbors. Thus, when an agent’s stack of sand grains becomes too high, it transfers them to its neighbors, and settles back to a ground state of zero. During this step, no more grains of sand are randomly dropped on the table.



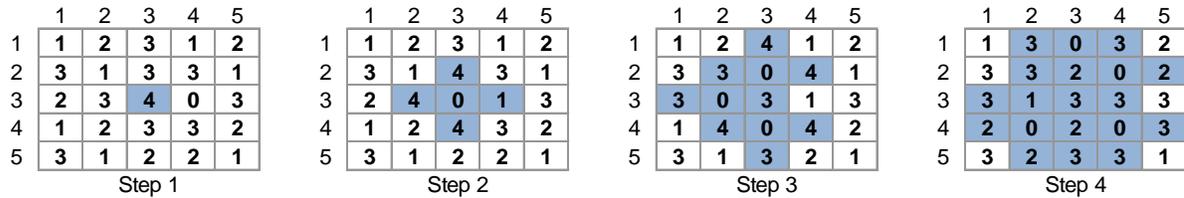
The previous step is repeated until there are no more agents with a state of 4 (ie, with four grains of sand).

After all agents with four grains of sand have been reset to zero, a grain is again randomly dropped on the table and the process recommences.

C. CA MODELS CONTINUED

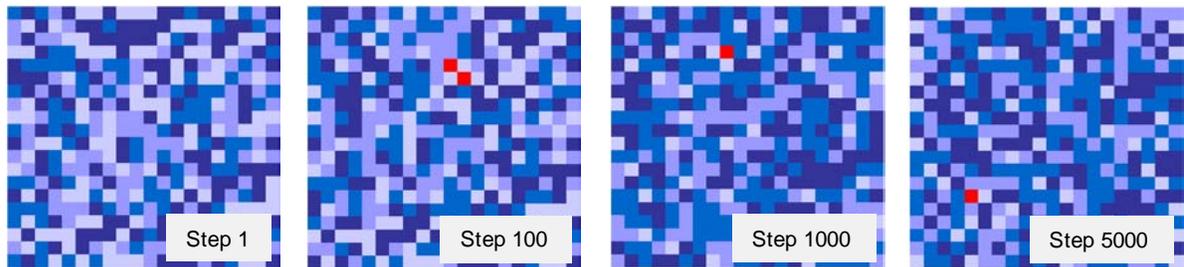
3. Sandpile continued

A key characteristic of the Sandpile model is an ‘avalanche’. The following sequence of model steps shows the progress of an avalanche:



At step 1, a grain of sand is added to the central agent, making it critical. In step 2, it topples, but makes three other agents critical. This domino, or avalanche, effect continues to step 4, at which no agents are critical, and another grain can again be randomly added. The shaded areas in the sequence trace the avalanche’s impact. This avalanche lasted four time steps and involved eight critical agents.

The following sequence shows the evolution of the Sandpile model with 400 agents over 5000 time steps, after starting from a random distribution of states 0, 1, 2, and 3 among the agents.²¹ State 0 corresponds to the color white, states 1-3 are successively darker hues of blue, and state 4 is red.



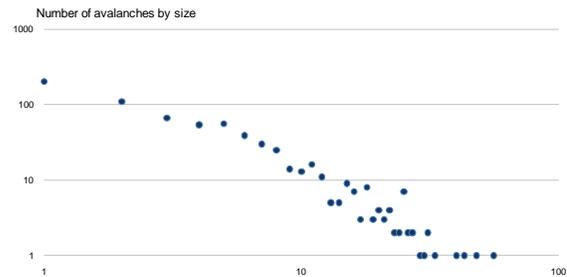
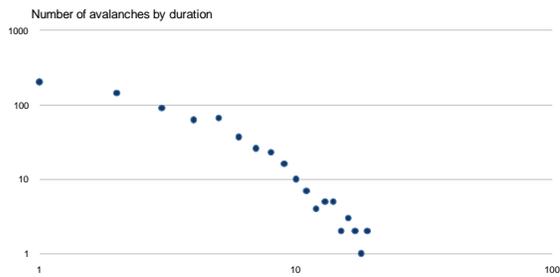
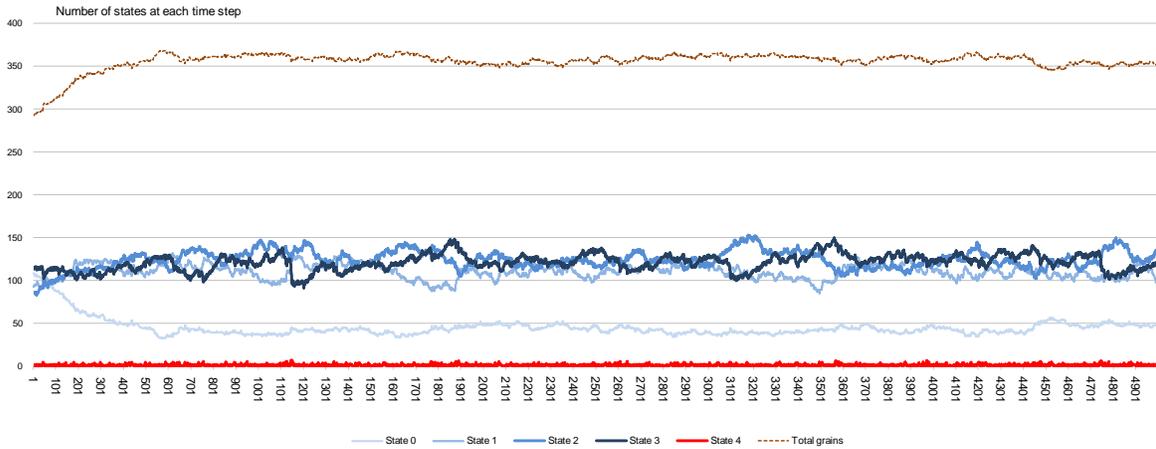
As you can see, over the 5000 steps, the agents are subtly organizing themselves: at step 5000 there are fewer contiguous agents in state 0, and the pattern is beginning to look more like a quilt than a random collection of colors. This is an example of Stuart Kauffman’s ‘order for free’, also known as ‘spontaneous order’ and ‘self-organization’ (see Chapter one).

²¹ Created with the Excel Sandpile model, with the following parameters: number of rows: 20; number of columns: 20; number of time steps: 5000; initialization option: “Randomly distributed 0, 1, 2, and 3”; random number seed: 3.

C. CA MODELS CONTINUED

3. Sandpile continued

The following charts analyze the 5000 steps of the Sandpile simulation:



The bottom two log-log charts demonstrate that the distribution of the model's avalanches are close to a power law. (Interestingly, one of the avalanches involved 56 agents at state 4, a result that in a normally-distributed world of avalanches – governed by the laws of dice – simply would not happen.) Per Bak concluded that the model is therefore at, or at least close to, a state of criticality, a place between perfect randomness and perfect structure, a place some have called the 'edge of chaos'.

The point to emphasize is that the model achieved this state on its own, without outside manipulation or fine-tuning.

C. CA MODELS CONTINUED

3. Sandpile continued

But does the Sandpile correspond to reality? In 1995, researchers at the University of Oslo carefully investigated, and video-taped, Bak's Sandpile model using piles of long-grain rice (grains of sand were too sticky). The experiment lasted for more than a year, and involved five scientists. They found that the rice pile behavior conforms closely to the Sandpile model. And, on a grander scale, sand slides in the Nepal Himalaya mountains have also been shown to conform to the Sandpile model.²²

The Sandpile model also suggests a reason why we find the power law ubiquitously in nature and in social systems: Just as a steady trickle of sand drives a sand pile to organize itself into a critical state, a steady input of energy, work, or money drives many systems to organize themselves to a state of criticality – with avalanches rearranging their structure just often enough to keep them poised on the edge. Thus, self-organized criticality may be one of the mechanisms that give rise to complex systems.

Per Bak also claims that his model applies to sediment deposits, earthquakes, solar flares, the demise of dinosaurs, the brain, traffic jams, and even social economies (more about this in Section D – Practical applications, below).

But self-organized criticality and the Sandpile model only describe the mechanism and overall statistics of avalanches in aggregate; they say nothing about individual avalanches. In fact, a corollary of self-organized criticality is that large avalanches – catastrophes – within a critical complex system are unpredictable.

As we conclude this exploration of the Sandpile model, it may interest you that Per Bak demonstrated that the Game of Life operates at or near a state of criticality, but not self-organized criticality. Amazingly, after examining millions of other simple 2D-CA behavior rules with Moore neighborhoods like the Game of Life, scientists did not find even one other that was critical. The Game of Life is so exceedingly rare that it cannot be self-organized; John Conway worked hard to fine-tune it.²³

²² Accounts of these two experiments are in Bak (1996), pages 69-78.

²³ Bak (1996), pages 105-112.

C. CA MODELS CONTINUED

4. Forest fire

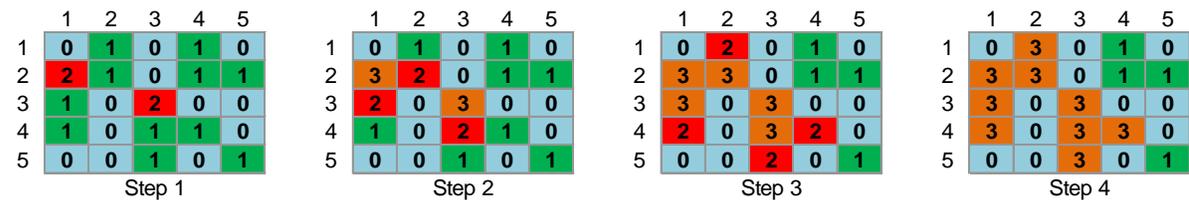
In 1990, Per Bak introduced another model of self-organized criticality, the Forest fire model. The Forest fire model is a 2D-CA on a lattice with a periodic boundary. Each agent represents a plot of forest ground, and has four states:

-  **State 0:** a plot without a tree (represented by blue)
-  **State 1:** a plot with a fully-grown tree
-  **State 2:** a tree on fire
-  **State 3:** a burned plot

The Forest fire model’s behavior rule is first order with radius one, and applies to von Neumann neighbors:

1. Fully grown trees (state 1) are initially randomly distributed according to a fixed probability.
2. Trees initially on fire (state 2) are also randomly distributed according to a fixed probability.
3. At each time step after the first, a tree can grow from an empty plot with a fixed probability. A tree sprouts and becomes fully grown (state 1) in one time step.
4. At each time step after the first, lightening can set a tree on fire (state 2) with a fixed probability. (This feature was not in Bak’s original model. It was added later to allow the model to evolve to criticality.)
5. A fire spreads to its nearest neighboring trees (in a von Neumann neighborhood) at the next time step.
6. A tree burns down completely in one time step and becomes a burned plot in the next time step (state 3).

For example, the following sequence shows the evolution of two fires until they burn out.

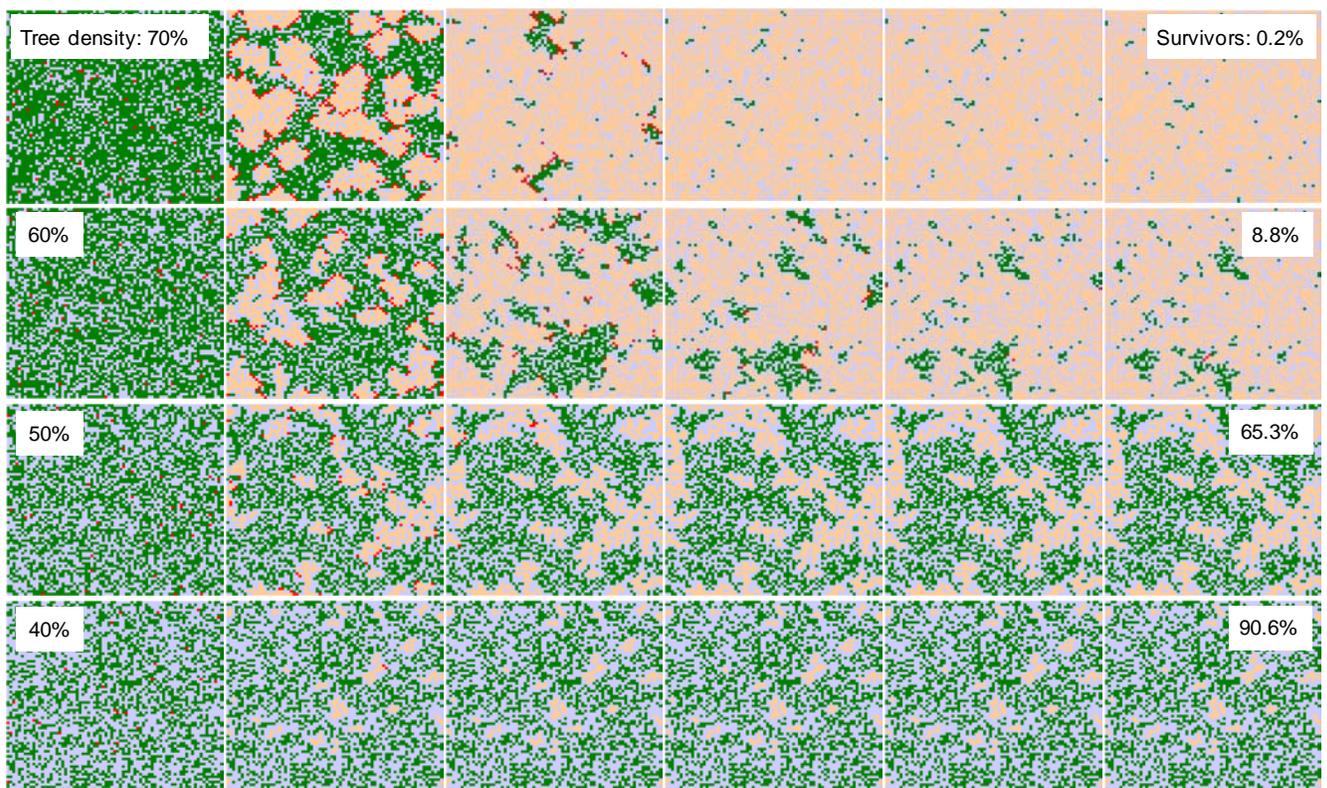


The sequence is simplified, because at steps 2-4, a new tree could grow, or lightening could strike a tree and set a new fire.

C. CA MODELS CONTINUED

4. Forest fire continued

Each row of the following shows Forest fire model results at time steps 1, 10, 20, 30, 40, and 50 for a particular initial tree density. For example, the first row shows the results for an initial density of 70 percent. The last frame of each sequence gives the number of trees that survived the fire. (eg, for the fire in the first row, only 0.2 percent survived). For each set of results, one percent of the trees are on fire at step one, and after step one there is no new tree growth and no lightning to set new fires.²⁴



As the population of trees becomes more dense, a small initial fire will have more devastating consequences. Once again we see the structure of agent relationships has a dramatic impact on a system's behavior.

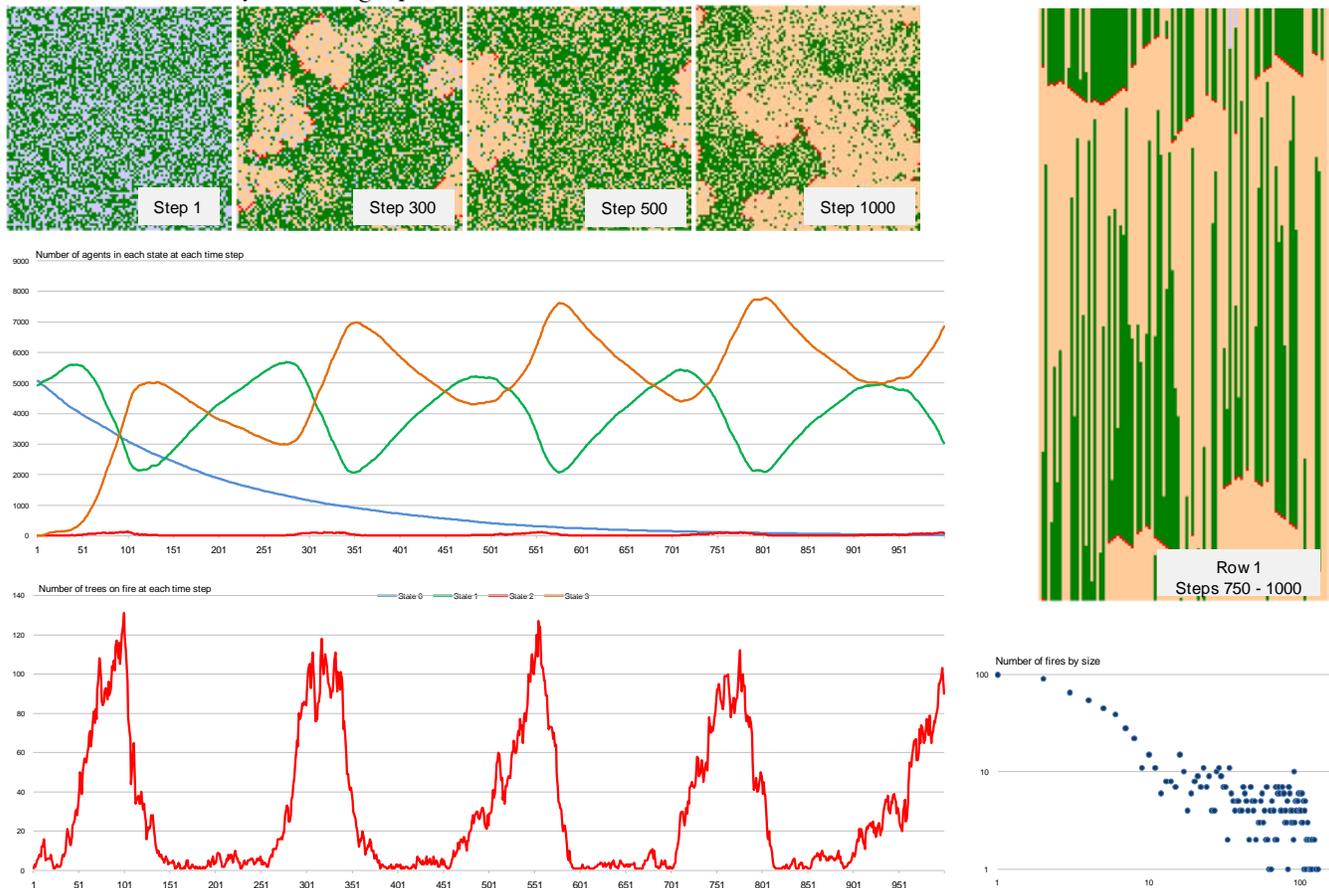
²⁴ Created with the Excel Forest fire model, with the following parameters: rows: 75; columns: 75; number of time steps: 50; growth probabilities: initial: .40, .50, .60, .70; on unburned plot: 0.0; on burned plot: 0.0; fire probabilities: initial: 0.01; subsequent: 0.0; action if no fire: "Let it remain burned out"; random number seed: 2.

C. CA MODELS CONTINUED

4. Forest fire continued

Following are remarkable results for 1000 steps of a Forest fire model with a growth probability of 0.005 on burned and unburned forest plots, and a lightning probability of 0.00001.²⁵ The four sequential diagrams show the agents’ states at times 1, 300, 500, and 1000. Their patterns look like fluid turbulence. (In fact, Per Bak’s initial paper introducing the Forest fire model was titled “A forest-fire model and some thoughts on turbulence”.)

As the two line plots show, the number of trees, burned plots, and fires all exhibit periodic variation with steep slopes. Perhaps most remarkable, the log-log scatter plot (lower right) shows that the system has nearly self-organized into a critical state, with the number of fires nearly following a power law.



²⁵ Created with the Excel Forest fire model, with the following parameters: rows: 100; columns: 100; number of time steps: 1000; growth probabilities: initial: .50; on unburned plot: 0.005; on burned plot: 0.005; fire probabilities: initial: 0.0005; subsequent: 0.00001; action if no fire: “Start one new random fire.”; random number seed: 2.

C. CA MODELS CONTINUED

5. Other CA

Beyond the relatively simple models introduced in this section, the CA concept can be broadly extended:

- **The number of agents and agent states can be expanded.** For example, von Neumann's original proof of the CA universal computer employed 29 states per agent, and about 200,000 agents.
- **The behavior rules can be more complex.** They can be of higher order (incorporating information from periods earlier than the last time step); they can extend past immediate neighbors with longer radii; and they can even involve learning and adaptation, subjects we explored in Chapter two and will encounter again in the next chapter.
- **The graph can be more realistic.** Although 1D and 2D lattices simplify computations and visualization, they are not particularly realistic. As we saw in the network chapter, real networks, and particularly the social networks with which actuaries work, usually have more interesting graphs. Many social networks have agents that connect with more than an average number of other agents, and most social networks have agents with a few connections reaching beyond a local neighborhood.

In the next section (practical applications) you will learn many ways to extend the CA concept.

D. PRACTICAL APPLICATIONS

As with networks, it appears that in actuarial work, cellular automata are virtually unknown. In the literature search for this report, I looked in the major actuarial databases for resources with either of the terms ‘cellular automata’ or ‘cellular automaton’. There was not a single relevant hit.²⁷

Cellular automata have been applied to many other disciplines. For example, in *A new kind of science*, Wolfram shows that 1D-CA and 2D-CA help explain many phenomena that were previously unexplained or misunderstood, including the growth of snowflakes and crystals, fracture patterns of solid materials such as rocks, fluid flow, growth of plants and animals, biological pigmentation patterns, financial markets, the limitations of mathematics, irreversibility and the Second Law of thermodynamics, the discrete nature of the universe, gravity, space, and time.²⁸

Starting with modeling of financial market behavior, this section will describe eight applications of CA that actuaries may find interesting and useful. Many of these applications employ advanced techniques that will expand your understanding of what CA models can do.

1. Financial market behavior

Why do the prices of financial instruments (stocks, bonds, currencies, etc.) often fluctuate with apparent randomness, and sometimes with catastrophic swings? Per Bak and Stephen Wolfram offer two similar explanations.

Bak uses his Sandpile model to suggest that such fluctuations are inevitable and perhaps even desirable (see sidebar). And Wolfram suggests that simple 1D-CA provide an explanation. To model a financial market, he suggests using a 1D-CA where each agent corresponds to a single trading entity, and its two states are ‘buy’ and ‘sell’. The behavior rule (how it decides whether to buy or sell) is based on the states of its immediate neighbors in the previous time step:



²⁶ Bak (1996), pages 191-198

²⁷ See the section titled *Finding the essential resources* in this report

²⁸ Wolfram (2002), pages 363-546.

Sandpiles and economics

Per Bak concludes his book *How nature works* with the following observations about economics:

“Our conclusion is that the large fluctuations observed in economics indicate an economy operating at the self-organized critical state, in which minor shocks can lead to avalanches of all sizes, just like earthquakes. The fluctuations are unavoidable. There is no way that one can stabilize the economy and get rid of the fluctuations through regulations of interest rates or other measures. Eventually something different and quite unexpected will upset any carefully architected balance, and there will be a major avalanche somewhere else in the system. ...”²⁶

If economics is indeed organizing itself to a critical state, it is not even in principle possible to suppress fluctuations. Of course, if absolutely everything is decided centrally, fluctuations could be suppressed. In the sandpile model, one can carefully build the sandpile to the point where all the heights are at their maximum value, $Z=3$. However, the amount of computations and decisions that have to be done would be astronomical and impossible to implement. And, more important, if one indeed succeeded in building this maximally steep pile, then any tiny impact anywhere would cause an enormous collapse. The Soviet empire eventually collapsed in a mega-avalanche (not predicted by Marx). But maybe, ... the most efficient state of the economy is one with fluctuations of all sizes. ...”

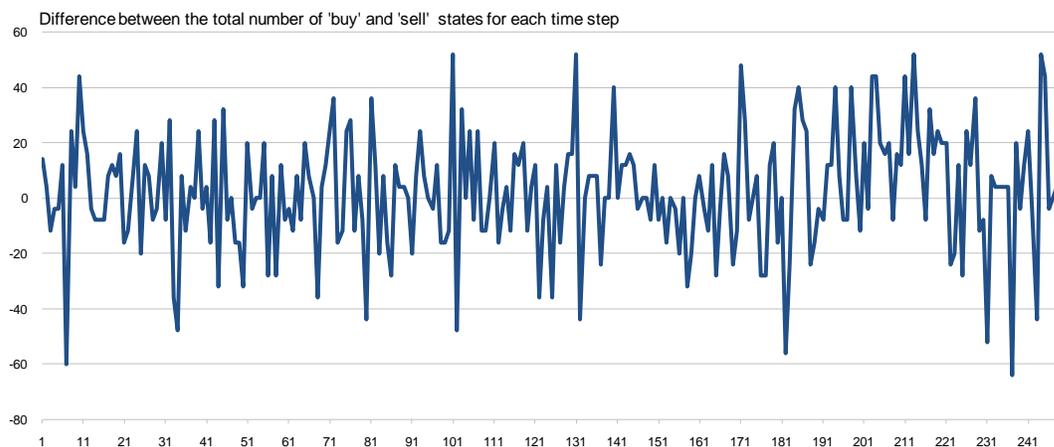
Maybe Greenspan and Marx are wrong. The most robust state for an economy could be the decentralized self-organized critical state of capitalistic economics, with fluctuations of all sizes and durations. The fluctuations of prices and economic activity may be a nuisance (in particular if it hits you), but that is the best we can do.”²⁶

D. PRACTICAL APPLICATIONS CONTINUED

1. Financial market behavior continued

Thus, when both neighbors either bought or sold, the agent sells; and when only one neighbor bought, the agent buys.

Wolfram further suggests that, for one time step, the difference between the total number of buy states and the total number of sell states provides an adequate proxy for the market price. Over the model's many time steps, this difference represents the market's price movement. The plot below shows such movement, over 250 time steps for such a 1D-CA with random initial assignment of buy and sell states:



Its movement is random, like the apparent random movement of real financial markets. In fact, the plot is one we explored earlier, because the behavior rule of Wolfram's market model is simply Rule 90 (in disguise).²⁹

In 2003, following Wolfram's idea, Yi-ming Wei et al developed a 2D-CA model of investment behavior in a stock market. In their model, each agent represents a stock market investor. Each agent has three states: 'buy', 'sell', and 'hold'. The behavior rule depends on the states of the agent and its Moore neighbors in the previous time step, an 'imitation' or 'contrarian' probability for each agent, and a system-wide variable representing positive or negative information about stocks.

²⁹ See Wolfram (2002), pages 429-432.

D. PRACTICAL APPLICATIONS CONTINUED

1. Financial market behavior continued

The authors conclude:

- The evolution of the stock market cannot be predicted.
- Stronger investor imitative behavior corresponds to greater market instability.
- “According to these primitive results of simulation of investment behavior in the stock market, we can see that cellular automata, an important tool for probing the complexity of systems, manifest huge potential in the research on the complexity in stock markets.”³⁰

2. Disease propagation

Recognizing the limitations of conventional epidemiological models, in 1996 Anders Johansen employed a Forest fire model (with lightening) to simulate disease propagation. His model is different from the classic Forest fire model in two respects:

- the fire (disease) spreads to neighbors with a fixed probability (whereas, in the classic model, the probability is unity)
- each agent has a fixed probability of being immune from the disease.

After discussing his model’s results, he writes, “In conclusion, it seems plausible from the simulations that a spatial organization of infectious individuals on fronts will give sustained temporal oscillations of infectious individuals no matter the specific manner in which new susceptible individuals are introduced. Furthermore, it seems likely that a sufficient condition for such a behavior is simply a n.n. (*ie, neighbor-to-neighbor*) interaction transmitting the disease. Given the extreme simplicity of the model, n.n. interactions and one parameter determining the (mean) input of susceptible individuals, we believe the model provides new insight into the sufficient conditions for obtaining recurrent epidemic behavior.”³¹

The fronts and oscillations that Johansen observed are the same as the areas of cyclic turbulence that we saw in the classic Forest fire model.

³⁰ Wei, Ying, Fan, & Wang (2003)

³¹ Johansen (1996)

D. PRACTICAL APPLICATIONS CONTINUED

2. Disease propagation continued

Building on Johansen's work, in their 1997 article, Rhodes et al demonstrate how the Forest fire model (again with lightening) accurately simulates outbreaks of measles, whooping cough, and mumps on the Faroe Islands (Danish islands for which there are long-running and detailed epidemiological records). Their model accurately captures the observed power-law and cyclical dynamics of the disease outbreaks, aspects that conventional epidemiology cannot capture.

They conclude, "Finally, we believe that this study illustrates how ideas developed to understand critical phenomena in the physical sciences can be used to gain insight into the dynamics of real epidemic diseases. It illustrates that biological systems can also exhibit critical phenomena, that the conditions for critical behavior to occur can arise quite naturally and that simple models which capture the essential interactions can explain the dynamical behavior of a number of different diseases in a human population."³²

3. Opinion dynamics

In 2006, Tony Wragg, member of a Command and Control Division of the Australian Department of Defense, developed a CA model to demonstrate the power of agent-based modeling to simulate how people form their opinions, and how well-deployed information can shape that formation.³³ His goal was to address the following questions:

- How well can an information campaign be represented by a computer simulation?
- What effect do the media have on a population's opinion?
- Does the population density of a population affect the level and rate of opinion change?

His model and its results have obvious implications for military propaganda campaigns, but, perhaps less obviously, are also relevant to marketing strategies for all types of organizations, including actuarial organizations.

³² Rhodes, Jensen, & Anderson (1997)

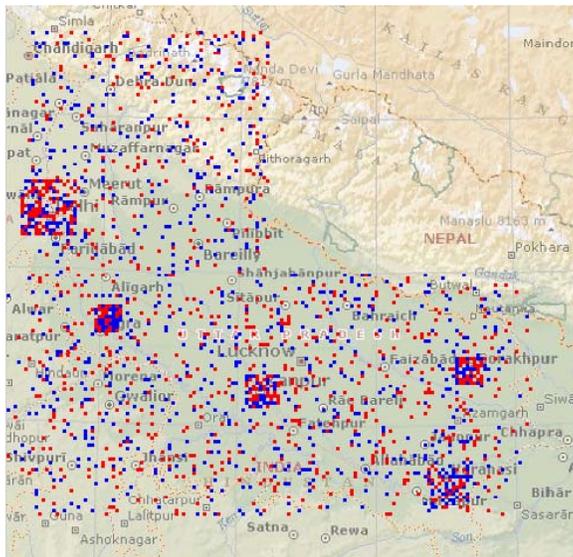
³³ Wragg (2006)

D. PRACTICAL APPLICATIONS CONTINUED

3. Opinion dynamics continued

His model simulates opinion dynamics among people of a northeastern province of India, Uttar Pradesh, concerning the issue of polio vaccination. Since 1988, the World Health Organization has tried to eradicate polio in India. The model uses data obtained as part of that effort.

The model is a 2D-CA with 150 columns and 150 rows. Each agent has three states: empty (no family occupies it), a family whose opinion favors polio vaccination (blue), or a family whose opinion rejects vaccination (red). The following image shows the initial agents states, superimposed on a map of Uttar Pradesh. Many of the agent states are empty, because that part of the grid lies over the neighboring country of Nepal.



The relative distribution of opinions and clustering of agents are based on actual data.

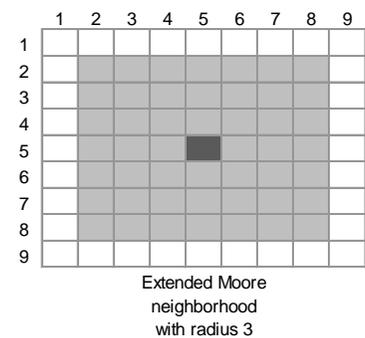
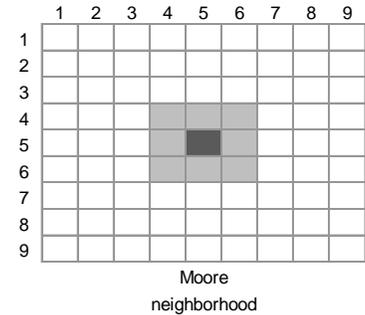
In addition to its state, each agent has the following attributes:

- Social influence = S (a normally distributed value, with mean 5 and variance 4)
- Resistance to change = B (a constant, equal to 2)
- Susceptibility to media influence S_M (a normally distributed value, with mean 5 and variance 4)

D. PRACTICAL APPLICATIONS CONTINUED

3. Opinion dynamics continued

Each agent’s behavior rule determines whether it has an opinion supporting (blue) or rejecting (red) vaccination. The rule depends on the state and other attributes of its neighbors in a Moore neighborhood with radius 32. The graphics at right show two Moore neighborhoods, one with a radius of 1, and one with a radius of 3. So, the neighborhood of radius 32 that influences an agent is quite large. However, the greater distance a neighbor is from the central agent, the less its impact.



When an agent is influenced by other agents, a theory called ‘social impact theory’ asserts that the level of social influence experienced by the agent is represented by:³⁴

$$I_i = -S_i B - O_i O_M S_{M_i} - \sum_{j=1, j \neq i}^N \frac{S_i O_j O_i}{d_{i,j}^\alpha}$$

where:

- I_i is the amount of social pressure upon agent i ($-\infty < I_i < \infty$)
- O_i is the agent i ’s opinion. It takes the values 1 (supporting vaccination) and -1 (rejecting)
- O_M is the opinion of the media (± 1)
- S_i is agent i ’s strength or influence ($S_i > 0$)
- S_{M_i} is the influence of the media on i ($S_{M_i} > 0$)
- B is agent i ’s resistance to change ($B > 0$; for this model 2)
- d_{ij} incorporates both the grid distance and religious distance between agents i and j ($d_{ij} \geq 1$)
- α is the distance decay exponent ($\alpha \geq 2$, for this model 2)
- N is the total number of agents in i ’s neighborhood of influence

Each agent’s behavior rule is then given by:

$$O_i(t + 1) = \begin{cases} O_i(t) \text{ with probability } \frac{\exp(-\frac{I_i}{T})}{\exp(-\frac{I_i}{T}) + \exp(\frac{I_i}{T})} \\ -O_i(t) \text{ with probability } \frac{\exp(\frac{I_i}{T})}{\exp(-\frac{I_i}{T}) + \exp(\frac{I_i}{T})} \end{cases}$$

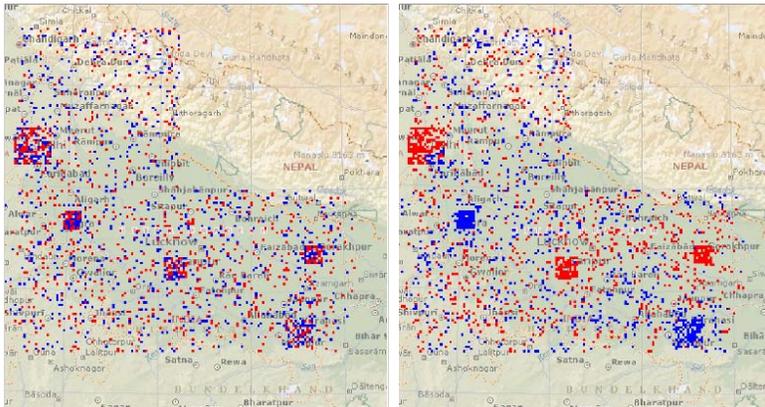
where T is a variable representing the volatility of individual decision making. Higher T reduces the probability that an agent will change its opinion for a given level of social influence.

³⁴ For more information about this approach, see Wragg (2006), Nowak & Lewenstein (1996), and Sobkowicz (2003).

D. PRACTICAL APPLICATIONS CONTINUED

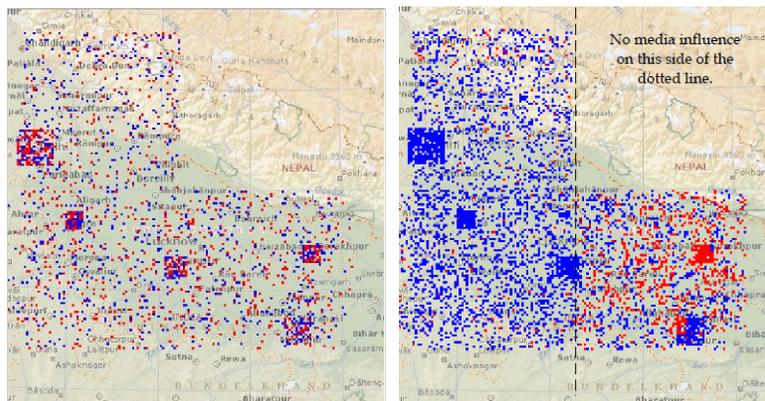
3. Opinion dynamics continued

The images below show the impact of not having any mass media messages about polio vaccination. The first image shows an initial opinion distribution with 50 percent of families supporting vaccination, and the second shows the opinion distribution, without mass media messages, after reaching equilibrium (after about 200 time steps).



At equilibrium, it is clear that like opinions tend to cluster, especially in areas of high population density.

The images below show the impact of mass media. The first image shows an initial 50 percent opinion distribution, and the second shows the opinion distribution at equilibrium after delivery of mass media messages to only part of the population.



Clearly, media coverage has a profound impact on opinion dynamics, particularly in areas of high population density.

Conclusions

The author of the simulation study found that the model demonstrated the typical characteristics of real-world opinion dynamics, namely similar-opinion clustering, polarization, and non-linearity of opinion change over time. Thus, he concluded that an information campaign could be represented by an agent-based computer simulation.

His model also provided useful answers to the author's initial questions about the impact of media and population density on opinion dynamics.

He made the following useful comments about such models:

- Validation of simulation results is challenging, because of the difficulty of running controlled experiments with large populations.
- “Obtaining the requisite data to enable the models to accurately represent a given society represents the biggest hurdle in social simulation today. Nevertheless, simulation always has a valuable role in helping to clarify ideas and theories even if complete validation cannot be carried out.”
- “Agent-based computer simulations have become the most powerful tool in studying the dynamics of social theories. . . . Agent-based simulations enable the dynamic and emergent properties of social influence campaigns, such as polarization and clustering, to be reproduced and analyzed. Although the phenomenon of social change is very complex, applying and extending theories such as the theory of social impact enabled the most critical factors of social influence to be isolated and varied systematically within a very simple model.”

D. PRACTICAL APPLICATIONS CONTINUED

4. Distribution of surgical volume

U.S. healthcare is largely inefficient and often ineffective. A case in point is complex surgical procedures. Extensive research over 30 years has conclusively demonstrated an inverse relationship between provider case volume and postoperative mortality and complications. That is, the more complex surgeries of a particular type (such as pancreatic resections) that a surgeon performs, the better the outcome. Common sense. Yet, because of the way the U.S. health system is structured, often there are many surgeons who perform complex surgeries at low volume. For example, in 2004 there were 1,000 pancreatic resections performed in Florida. About 300 surgeons and 85 hospitals contributed to this volume, and more than half of the surgeons performed fewer than 2 such procedures. Mortality and complication rates for these surgeons were 2-3 times higher than for their peers.

In 2008, James Studnicki et al wanted to study the potential impact of informed patient choice on this state of affairs.³⁵ They developed a CA model with the following characteristics:

- **Agent types:** Patients, hospitals, surgeons, and payers (insurance companies). Only patients have behavior rules to make decisions.
- **Patient attributes:** Geographical location. Each patient has a condition that requires complex surgery.
- **Hospital attributes:** Geographical location, cumulative patient satisfaction with the hospital.
- **Surgeon attributes:** Age (normal distribution with mean 47.5 and standard deviation 5.0, constrained to be between age 30 and 60; at age 60 a surgeon is removed), geographical location, number of procedures performed, cumulative patient satisfaction with the surgeon.
- **Relationships:** Each patient subscribes to one payer. Hospitals are related to surgeons in a random network generated from an input parameter for the network's density, and payers are related to hospitals in a similarly-generated random network.
- **Time step:** The number of days represented by each time step is an input parameter.

³⁵ Studnicki, Eichelberger, & Fisher (2009).

D. PRACTICAL APPLICATIONS CONTINUED

4. Distribution of surgical volume continued

- **Patient behavior:** The behavior rule determines what hospital-surgeon pair the patient chooses to perform the surgical procedure. The patient evaluates each eligible pair in the payer’s network. The pair that has a maximum ‘fitness’ is selected. To be eligible, the surgeon must satisfy the following criteria:

Proximity: The surgeon must be located within a specified distance from the patient (an input parameter).

Workload: The surgeon must not have exceeded a given number of surgeries per time period (an input variable).

$$fitness = C^{W_C} \times E^{W_E} \times S^{W_S}$$

where:

$$C = \frac{1}{\log(1 + dist(patient, hospital) + dist(patient, surgeon))}$$

$dist(patient, x)$ is the patient’s travel distance.

$$E = \frac{\text{procedures}(surgeon) \wedge 10}{100}$$

The floor of 10 procedures per surgeon reflects the inability of patients to distinguish among low-volume surgeons.

$$S = \frac{1}{1 + e^{-(0.8 \times satisfaction(surgeon) + 0.2 \times satisfaction(hospital))}}$$

Each time a surgery is performed, the satisfaction score is updated:

$$satisfaction(t + 1) = satisfaction(t) + outcome(t)$$

$$\text{where } outcome = \begin{cases} 0.5 & \text{if complication} \\ 0.1 & \text{if death} \\ 1.0 & \text{otherwise} \end{cases}$$

The weights W_C , W_E , and W_S are input parameters.

Output: Number of high-volume surgeons (surgeons who performed more surgeries than average), average surgeon age, procedures per year per surgeon, death and complication rates (in total and per surgeon).

D. PRACTICAL APPLICATIONS CONTINUED

4. Distribution of surgical volume continued

At each time step, the model:

- Randomly generates new patients requiring surgery.
- Invokes each patient's behavior rule to select a surgeon and hospital.
- Determines the surgical outcome based on outcome rates built into the model.
- Increments variables to reflect the outcome, and increments the time step.

Each simulation was run for 10,000 time steps.

The model produced several interesting results. For example, according to the model, a payer that allows patients to select only surgeons and hospitals from within its network actually increases the number of expected complications. At first this result may seem counter-intuitive, but the reason is that networks can restrict the number of cases that potentially high-volume surgeons can perform, thus limiting their experience and increasing their rate of negative outcomes.

The authors conclude, "This exploratory model suggests multi-agent simulation methods can be helpful in understanding the complex interactions which are operative within the U.S. healthcare industry. We have focused upon the relationship involved in the performance of complex surgeries, especially those for which there is a significant likelihood of an adverse outcome in the form of a post surgical complication, or even death. Our model activated only the patient agent and determined passive roles for other agents. Future developments will involve activation of the other agents. Surgeons, for example, are likely revenue maximizers who determine the composition of their surgical caseload based, at least partially, upon the revenue received from each case. In that context, various 'complexity mixes' will result in a range of incomes. Modeling surgeon choice based upon revenue, workload, convenience, career phase and other factors will enable a more valid portrait of patient/surgeon interaction. Similarly, the hospital is interested in revenue maximization consistent with the best outcomes, i.e., minimizing deaths and complications. Since hospitals offer administrative privileges to physicians, they should be interested in minimizing the number of low volume surgeons who hold privileges."

D. PRACTICAL APPLICATIONS CONTINUED

5. Retirement incidence

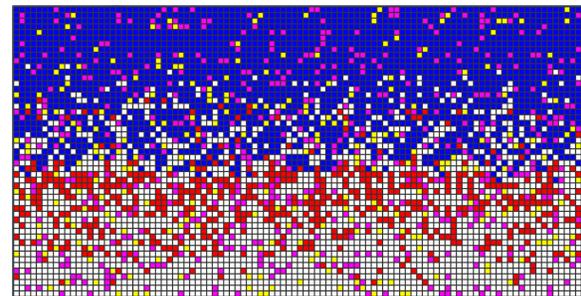
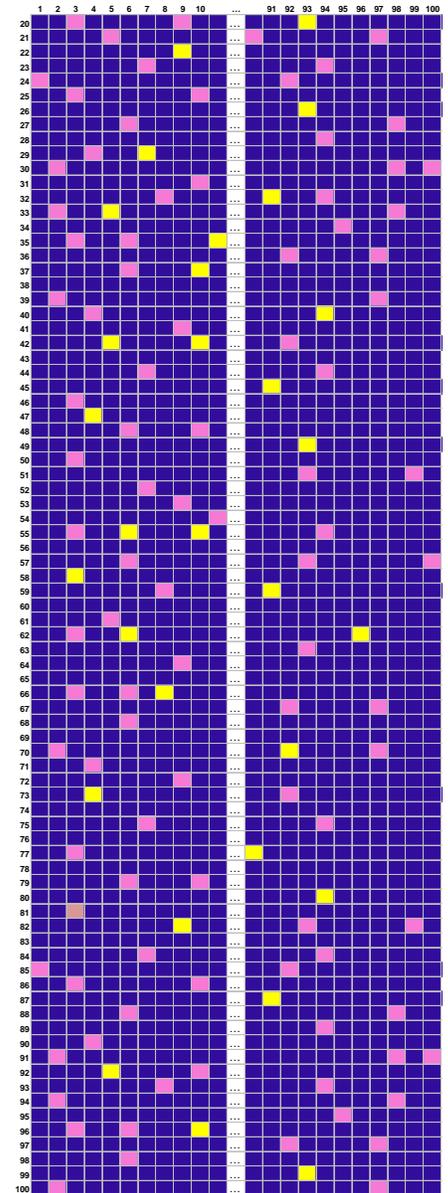
In 1961, the U.S. Congress reduced the minimum Social Security retirement age from 65 to 62. Yet, it took many more years than originally anticipated for the average retirement age to approach 62. To policy makers and economists alike, this result was puzzling.

Robert Axtell and Joshua Epstein created a CA model to solve the puzzle. The model is a 2D-CA, but with several interesting modifications that differentiate it from the classic simple 2D-CA.³⁶ Their 2D grid is 100 columns by 81 rows, representing 8,100 agents. All the agents in one row are in one age group. Thus, there are 81 annual age groups, from age 20 to age 100. As the simulation progresses, each time step represents one year, and the agents in one row at one time step move to the next row at the next time step. Also, at each new time step, the previous age-100 cohort is removed from the grid, and a new age-20 cohort is added.

Each agent has three attributes related to how it makes its retirement decision: ‘rational’, ‘random’, or ‘imitative’. Rational agents all retire at the earliest possible age allowed by government policy (this age is arbitrary, but for the model, it is assumed to be 65). Random agents retire at any eligible age, with probability 0.5. Imitator agents base their decision on the decisions of their social network (explained more fully below). On the grid, rational agents are represented by the color pink, random agents by yellow, and imitator agents by blue (see the grid to the right).

Each agent also has three states: ‘working’, ‘retired’, or ‘dead’. Retired agents are represented by the color red, and dead agents by the color white. In the model’s simulation runs, these colors are laid over the colors representing the decision types, and only the rows for ages 55 to 100 are displayed (see the grid below right).

The key to the model is that an imitator’s behavior rule for deciding to retire is based on the fraction f of agents in its social network who have already retired. The imitator agents have a randomly assigned ‘imitation threshold’ t ; if at any time $f \geq t$, the agent retires.



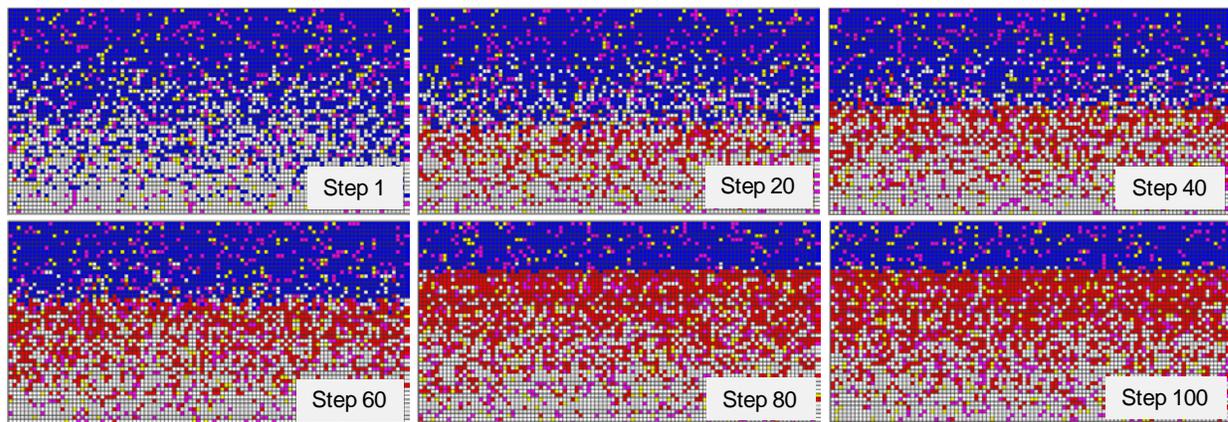
³⁶ Epstein (2006), Chapter 7: Timing of retirement. This book is one of this report’s Top ten Complexity Science books.

D. PRACTICAL APPLICATIONS CONTINUED

5. Retirement incidence continued

Rather than base each agent's social network on a classic von Neumann or Moore neighborhood, Axtell and Epstein assign each agent its own network. For each agent, the number of other agents in its network is a uniform random number between the two parameters 'minimum network size' and 'maximum network size'. For example, these two parameters might be 10 and 25. Then, the model randomly selects a number of agents between 10 and 25, within an age range of ± 5 years. Thus, in its network one agent age 60 might have 15 other agents ranging in age from 55 to 65, while another agent age 60 might have 25 ranging in age from 59 to 63.

Following is a sequence of model results for a minimum retirement age of 62, 10 percent rational agents, 5 percent random agents, and 85 percent imitative agents. The model demonstrates that it can take a long time for agents to adopt a legislated retirement age.³⁷



As the authors point out, aside from offering a solution to the puzzle of slow adoption, this result is interesting because, even though the vast majority of agents are not rational (as in real life), the system as a whole reaches a rational decision. Thus, aggregate rationality arises from a system comprised of mostly boundedly rational agents, a surprising emergent property.

Perhaps more interesting, if the percent of rational agents is reduced to 5 percent, the time to arrive at a rational decision is much longer, and the route to aggregate rationality is not a smooth progression.

³⁷ This sequence was created using the model supplied on CD with Epstein (2006). Model settings were: Fraction random: 0.05; Fraction rational: 0.1; Min network size: 10; Max network size: 25; Retirement eligibility age: 62.

D. PRACTICAL APPLICATIONS CONTINUED

5. Retirement incidence continued

In concluding, the authors write, “While we have interpreted this model as applying to retirement, it could be applied to a wide range of settings in which social interactions mediate purely rational behavior. Obvious candidates include contagion behavior in markets, migration to different health plans, or the diffusion of technological innovations. In reality, these phenomena occur in social networks, while most existing models treat them either as occurring in ‘perfectly mixed’ environments or via local interactions on regular lattices or other highly specialized topologies. The agent-based computational approach is well suited to studying such processes with any topology of interactions.”

6. Policyholder lapse behavior

In the late 1990s, Ernst & Young created a strategic alliance with a company called the BiosGroup headed by Stuart Kauffman (see sidebar). One of their joint projects was to build a model of policyholder lapse behavior for guaranteed income variable annuity products, using Complexity Science concepts. The motivation for developing such a model was that such variable annuity products were just being introduced. So, there was no real-world experience to draw on for pricing, making statistical modeling techniques inapplicable.³⁸

They constructed a proof-of-concept CA model, with the following characteristics:

- **Agent types:** policyholders (customers) and brokers (sellers). The model can simulate millions of customers and thousands of sellers.
- **Customer attributes:** Age (uniformly distributed from 55 to 65), gender (randomly assigned), annuity type (uniformly distributed between 50 percent and 100 percent variable), and policy duration (each agent holds one policy, which can have any duration).
- **Relationships:** Each agent is related to one broker (in a skewed distribution, with the top 50 percent of brokers controlling 75 percent of the business), and to a group of ‘friends’ that it may canvass to make a lapse decision.
- **Time step:** Each time step is one year. Each customer makes a lapse decision every year.

³⁸ See Shumrak, Greenbaum, Darley, & Axtell (1999) and Shumrak & Darley (1999).

Santa Fe Institute progeny

The Santa Fe Institute has spawned so many business-oriented organizations that the area around Santa Fe is now known as the Info Mesa. Among its progeny are:

BiosGroup was founded in 1997 by Stuart Kauffman. Its goal was to commercialize Complexity Science software to help companies manage projects and supply chains. Its clients included Southwest Airlines, P&G, Ford, Boeing, Texas Instruments, and the Internal Revenue Service. At its peak, it employed about 150 people in offices in Santa Fe, Boston, London, Bulgaria, and Washington DC. In 2003, its consulting operations were acquired by NuTech Solutions.

Complexica was founded in 2001 by Roger Jones and John Casti. Its spinoff **Assurtech, Inc.** developed the model called *Insurance World* that we’ll explore in Chapter seven. In 2004, both companies were absorbed by ComodiCast Inc.

Prediction Company was founded in 1991 by Doyme Farmer and Norman Packard. It builds advanced financial market trading systems that incorporate Complexity Science-inspired forecasting techniques. In 2005, it was purchased by Union Bank of Switzerland.

D. PRACTICAL APPLICATIONS CONTINUED

6. Policyholder lapse behavior continued

- **Customer behavior:** The customer behavior rule determines whether a policyholder agent will lapse its annuity policy. A certain percentage (1 percent in the model) will die. Another percentage (2 percent) will lapse for random reasons, such as needing money to build a house. Another group receives advice from their sellers and will make a rational lapse decision by comparing the advice to the current financial performance of their annuity contracts. Another group will not receive any advice. 40 percent of these will make a rational lapse decision by comparing the performance of competitive products to the current financial performance of their annuity contract. The remaining group (the imitators) will base their decision on the decisions of their friends.
- **Seller behavior:** Sellers are motivated to move their customers into new products. Each time step, each seller sends out advice to a random percentage of its customers. The seller's advice is based on market scenarios, and provides an optimistic assessment of returns that customers could obtain from competitive products. The degree of optimism varies by broker. There are three market scenarios: a flat market, an inflationary market in which real returns are low, and a low-inflation market in which returns are good.
- **Output:** The model's primary output is annual lapse rates by policy duration. These results can be easily incorporated into actuarial pricing and risk analysis models.

In his article about the model, Michael Shumrak concludes, "...ABM (ie, agent-based modeling) provides a more robust approach to the real interaction of factors driving policyholder behavior. Effective use of these techniques can be applied to develop policy behavior dynamics for pricing product benefit designs, evaluating policyholder conservation programs and evaluating the impact of economic scenarios not seen before. Also, after constructing the behavior model, sensitivity tests can be performed to establish confidence intervals on the output. If we do so, we can immediately see to which of those parameters the new predictions are especially sensitive, and on which they depend weakly. This is tremendously important in situations for which there is no historical data."³⁹

³⁹ Shumrak, et al. (1999), pages 4 and 5.

D. PRACTICAL APPLICATIONS CONTINUED

6. Policyholder lapse behavior continued

In 2002, Charles Boucek and Thomas Conway, both actuaries, developed another CA model of policyholder lapse behavior. Their goal was to determine the impact that a rate change would have on policyholder retention and on a company's resulting profitability.⁴¹

Following are characteristics of their model:

- **Agent types:** Policyholders, the company changing its rates, competitors of the company, brokers. In order to conform to any company's book of business and its competitive environment, the number of agents is variable.
- **Policyholder attributes:** Age, gender, marital status, rating factors (such as driving record and vehicle usage), and policy factors (such as deductibles and liability limits).
- **Relationships:** each policyholder has a broker and a current insurance company.
- **Time step:** Each time step is one year. Each customer makes a renewal decision every year.
- **Policyholder behavior:** The behavior rule determines whether the policyholder will switch to a new insurance company or not. It is a two-tier decision: first, based on the premium increase from the insurance company, the policyholder decides whether to shop for new insurance; second, based on potential premium savings from competitors, the policyholder decides whether to switch. Associated with each level of premium increase and premium savings is a likelihood of a policyholder shopping and switching, respectively. The policyholder's decision is also influenced by its broker, but this behavior was not described.
- **Output:** The model tracks the distribution of policyholders across all rate classes before and after a rate change. It then uses this information to estimate total profitability and the volume of business that will be written.

The model is run for multiple iterations until its results converge to an equilibrium level of retention and profitability.

As with all the practical models in this report, the point of the model is to understand, not to predict.

⁴⁰ Boucek & Conway (2003), pages 162 and 165.

⁴¹ For a detailed description of the model, see Boucek & Conway (2003).

Why Complexity Science?

In justifying why they used a bottom-up Complexity Science approach for their model, Boucek and Conway write:

“Companies often have a number of ‘rules of thumb’ for determining the amount of a rate change that the market will bear, but very few rigorous models exist that attempt to estimate the likely customer reaction to a rate change. An approach to pricing that considers not only the impact of the new rates on the average premium charged, but also on the renewal behavior of policyholders can thus be a significant step forward for determining appropriate prices and likely future profitability. . . .

The current ‘rules of thumb’ approach may have been good enough at one time. It may also be true that this approach will be acceptable today in a situation where the rate change is simple. An example would be a rate change that only applies to the base rates. However, one of the trends for virtually all lines of business is that rate structures have become more refined over time. Using automobile insurance as an example, the number of different possible combinations of rate classes is so great that it is not possible to assess all of the changes that individual policyholders will experience in a rate change where base rates, territorial factors, driver classification factors and accident surcharges all change at the same time.

...
Another advantage of the ABM (ie, agent-based modeling) approach is that it allows for the modeling of emergent behavior. These are behavioral impacts, which may seem irrational at an individual level but are exhibited when the behavior of a group is analyzed as a whole. An example of this phenomenon is the observed behavior of groups of insured to leave when they are presented with a rate decrease.”⁴⁰

E. EXERCISES

1. For the class of 1D-CA, how many first-degree behavioral rules are there with a radius of one (rules that only relate to the states of next neighbors in the previous time step)? Now, after considering that a rule such as 10000011 is equivalent to the rule 11000001 (left-right symmetry), and that a rule such as 10100101 is equivalent to the rule 01011010 ('on'-'off' equivalence) how many truly unique behavior rules are there? It may help to put these rules in the Excel one-dimensional CA model to see their effects. Does the number of unique rules surprise you?

2. Consider CA-30 with one initial 'on' agent. Even though the sequence of states of the initial 'on' agent is perfectly deterministic, if you were simply given its state sequence from 1 to 200, and not given any other information, could you predict its state at time 300? Could you confidently predict how many '1's there would be between time 500 and 600? What does this say about our ability to predict such things as next quarter's health expenditure trend rate or next year's interest rate levels?

3. Per Bak's classic Sandpile model has 'cliff' boundaries. Implement a Sandpile model with periodic boundaries, and compare the resulting avalanche patterns with the two boundary types. Why do you think Per Bak preferred 'cliff' boundaries?

4. In the classic Sandpile model, sand grains are added to randomly-chosen cells (agents). Develop a Sandpile model that only adds sand grains to the center cell. Do results from this model differ from the classic model.

5. How would you create a one-dimensional Forest fire model? (Hint: you don't need to develop a new model to do this.)

6. To measure the randomness of simple 1D-CA and 2D-CA, the 1D and 2D models use a chart to compare the number of binary search strings of lengths 1, 2, 3, and 5 that are found in the agent's results, to the 80 percent/120 percent confidence interval of the expected number of such search strings if the agent's results were random. The models use a simplistic definition of confidence interval (see columns E and F of the sheet Analysis_SingleAgent). A more advanced approach would employ a binomial confidence interval reflecting p successes in n trials:

$$CI = n \times p \pm z_{1-\alpha/2} \times \frac{\sqrt{p(1-p)}}{n} \text{ with } z_{1-\alpha/2} = 0.84$$

Using this approach, would the results be materially different?

Complexity science – an introduction (and invitation) for actuaries

Chapter five: Cellular automata continued

F. TO LEARN MORE

To learn more about cellular automata, you may first enjoy watching Stephen Wolfram's presentation of *A new kind of science* at the University of California, San Diego.⁴² Then read the recommended pages of *A new kind of science*, one of this report's *Top ten Complexity Science books*.⁴³

Then, you may find it interesting to read Schiff (2008) and Barrat, et al. (2008).

G. REVIEW AND A LOOK AHEAD

This chapter introduced the second of the four archetypal Complexity Science models: cellular automata. CA models add behavior rules to the Network archetype, and so can trace the evolution of networks as component agents change their states based on behavior rules.

You learned about four types of cellular automata, 1D-CA, 2D-CA, the Sandpile, and the Forest fire; their terminology; how they are constructed and employed; and some examples of how they are applied. You also worked with four Excel CA models, and started to learn about agent-based modeling.

Next, we will explore a type of Complexity Science model that includes an environment. With an environment, agents can move around and interact with the environment. You will see that this model type becomes quite realistic.

There is a common misconception about CA, a mistake that even authors of CA books make. Many believe that an agent can move around on a 2D-CA grid, and so they classify models such as Schelling's segregation model (where agents move around on a grid) as CAs. This is not strictly correct. Because CA do not include an environment, CA agents cannot move. This misconception probably arose due to the appearance of motion in CA like The Game of Life. However, such motion is due to patterns of the changing states of stationary agents, not to agents who are moving. We will meet agents who can move in the next chapter.

⁴² Wolfram (2008), a YouTube.com video that lasts about 90 minutes.

⁴³ See the book's annotation in the section of this report titled Top ten complexity science books. You will find recommended sections for the book at the top of the annotation.

CHAPTER SIX: ARTIFICIAL SOCIETIES

Today's universities and think tanks are full of analysts who use multivariate equations to model the effects of changes in tax rates or welfare rules or gun laws or farm subsidies. I can easily envision a time, not long from now, when many of those same analysts will test policy changes not on paper but on artificial Americans that live and grow within computers all over the country, like so many bacterial cultures or fruit-fly populations. The rise and refinement of artificial societies is not going to be a magic mirror, but it promises some hope of seeing, however dimly, around the next corner.

Jonathan Rauch¹

A. INTRODUCTION

Joshua Epstein is a self-styled 'generativist', a colorful character (see sidebar), and largely responsible for changing the perspective of social science and economics from top-down to bottom-up.

Skipping high school, Epstein entered Amherst college to study piano and music composition. But there he fell in love with mathematics, and switched to study math and political economy instead. In 1981, he earned a PhD in political science from MIT, and soon thereafter began work at the Brookings Institution, Washington DC's oldest think tank.

In the early 1990's, he attended a conference at the Santa Fe Institute that changed his world view. Always enamored with models, at the conference he discovered models unlike any he had seen: models that grew lifelike artificial trees, flocks of birds, and schools of fish, from simple rules, from the bottom up. Inspired to try such models with human societies, Epstein returned to Brookings, and, in the cafeteria, told his colleague Robert Axtell about his idea. Together, on a napkin, they sketched such a rudimentary society, with agents moving around an artificial world, gathering its only resource – sugar. They called their artificial society Sugarscape, and spent the next few years working on it. In 1996, they published their paradigm-shattering book about Sugarscape, titled *Growing artificial societies*.³

This chapter introduces artificial societies (also called microworlds, surrogate worlds, virtual worlds, would-be worlds, sim-worlds, and even 'peasants under glass'), our third model archetype, which includes agents, agent relationships, agent behavior, and an environment. You will learn about two artificial society model types – Schelling's segregation model and Sugarscape – and how artificial society models are applied.

Epstein and Axtell

"Epstein is tall and portly, with a wild tuft of graying hair above each ear, a round face, and the sort of exuberant manner that brings to mind a Saint Patrick's Day parade more readily than a Washington think tank. 'No foam!' he roared, grinning to a Starbucks server one day when we went out for coffee. 'Keep your damn foam!'"

Anyone who notices Epstein is soon likely to encounter Robert Axtell, his collaborator and alter ego. A programming wizard with training in economics and public policy, Axtell is of medium height, quiet, and as understated as Epstein is boisterous. When he speaks, the words spill out so quickly and unemphatically that the listener must mentally insert spaces between them."

Jonathan Rauch²

Today, Epstein is the Director of the Brookings Institution Center on Social and Economic Dynamics, and a member of the External Faculty of the Santa Fe Institute. Axtell left Brookings in 2007 and is currently a professor at George Mason University.

¹ Rauch (2002), a fascinating article in The Atlantic Monthly about artificial societies, page 48.

² Rauch (2002)

³ Epstein & Axtell (1996), one of this report's *Top ten Complexity Science books*.

Complexity science – an introduction (and invitation) for actuaries

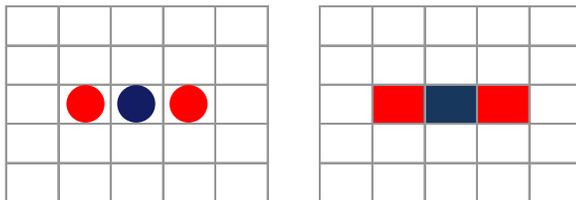
Chapter six: Artificial societies continued

B. BASICS

The distinguishing feature of our third model archetype, artificial societies, is an environment. An environment is a space on which agents move (such as a city), resources with which agents interact (such as groceries on shelves), or external events (such as the weather). With the addition of environments, the artificial society archetype closely resembles our real world.

In an artificial society model, an environment can be represented in many ways. It can be anything from a simple grid on which agents are located and move, to realistic 3D displays incorporating actual geographic coordinates and real-world features.

Although they may sometimes appear similar, artificial societies are quite different from our second model archetype, cellular automata (CA). To see this, consider the following two grids, where the grid on the left is a 2D artificial society model and the one on the right is a 2D-CA model:



In the CA model (on the right), because the grid represents the network of agent relationships, each of its cells is an agent. Thus, the CA grid includes 25 agents, with one agent in a blue state, two in a red state, and 22 in a white state.

By contrast, the grid on the left is an artificial society environment, with only three agents. The three agents can move around on the grid to any of the unoccupied 22 environment spaces.

To help distinguish the two model types, in this chapter, we will generally represent artificial society agents by round colored disks, rather than by colored cells.

C. ARTIFICIAL SOCIETY MODELS

There are two common artificial society model types that we will explore: the ‘Schelling segregation model’ and ‘Sugarscape’. This section describes each.

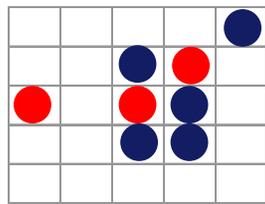
1. Schelling segregation model

Thomas Schelling is an American economist and political scientist who was awarded the 2005 Nobel Memorial Prize in Economic Sciences for enhancing our understanding of conflict and cooperation. One of his most famous results is the Schelling segregation model.

Developed in the 1960s and 1970s, the Schelling segregation model is one of the earliest Complexity Science models, with implications far beyond segregation. Because of his profound insight that a social system’s aggregate behavior may be quite different from what might be expected from extrapolating individual behavior, Thomas Schelling is sometimes called the father of agent-based modeling (see sidebar).

In the Schelling segregation model, the environment is a two-dimensional grid with periodic boundaries. An agent may be located in any cell of the grid, but each cell can contain only one agent. Agents have two states, blue and red, corresponding to the agents’ race (or, alternatively, its class, age, religion, language, sexual preference, income level, or any other distinction that may give rise to social segregation).

For example, the environment below contains eight agents, three red and five blue.



In this model, as you shall see, the network of relationships among agents is not a simple static lattice; rather it is complex and ever-changing, as in the real world.

Complicated social patterns from simple rules

Written in 2002 about Thomas Schelling:

“Today Schelling is eighty years old. He looks younger than his age and is still active as an academic economist, currently at the University of Maryland. He and his wife Alice live in a light-filled house in Bethesda, Maryland, where I went to see him one day not long ago. Schelling is of medium height and slender with a full head of iron-gray hair, big clear-framed eyeglasses, and a mild, soft-spoken manner. Unlike most other economists I’ve dealt with, Schelling customarily thinks about everyday questions of collective organization and disorganization, such as lunch-room seating and traffic jams. He tends to notice the ways in which complicated social patterns can emerge even when individual people are following very simple rules, and how those patterns can suddenly shift or even reverse as though of their own accord. Years ago, when he taught in a second-floor classroom at Harvard, he noticed that both of the building’s two narrow stairwells – one at the front of the building, the other at the rear were jammed during breaks with students laboriously jostling past one another in both directions. As an experiment, one day he asked his 10:00 A.M. class to begin taking the front stairway up and the back one down. ‘It took about three days,’ Schelling told me, ‘before the nine o’clock class learned you should always come up the front stairs and the eleven o’clock class always came down the back stairs’ – without, so far as Schelling knew, any explicit instruction from the ten o’clock class. ‘I think they just forced the accommodation by changing the traffic pattern,’ Schelling said.”

Jonathan Rauch⁴

⁴ Rauch (2002)

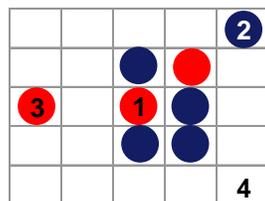
C. ARTIFICIAL SOCIETY MODELS CONTINUED

1. Schelling segregation model continued

Agent relationships

At any time step, an agent is related to other agents to which it is connected through contiguous locations on the environment. The definition of contiguous includes agents located in neighboring north, south, east, or west cells, as well as agents located in diagonally contiguous locations.

For example, agent 1 below is related to agent 2 (because they have a common immediate neighbor), but not to agent 3. And, of course, the environment location 4 is not an agent at all.



Agent behavior

An agent’s behavior rule determines whether it will move from one location on the environment to another. The rule depends on the agent’s preference for like-colored neighbors, quantified by a ‘neighbor preference percentage’. If the fraction of the agent’s like-colored immediate neighbors (within a Moore neighborhood of radius one) is greater than or equal to its neighbor preference percentage, the agent is happy and doesn’t move. However, if the fraction is less than the preference percentage, the dissatisfied agent moves to a random location that meets its preference threshold.⁶ The behavior rule is synchronously applied to all agents at every time step, and the time steps continue until all agents are satisfied and there is no more movement.

For example, if the neighbor preference percentage is 0.5, agent 1 in the diagram above would move, because it has 5 total immediate neighbors, but only 1 red neighbor. The fraction of its like-colored neighbors is $1/5 = .20$, which is less than 0.5.

Surprising aggregate results

Writing about the unexpected aggregate behavior of social systems, Thomas Schelling wrote:

“These situations, in which people’s behavior or people’s choices depend on the behavior or the choices of other people, are the ones that usually don’t permit any simple summation or extrapolation to the aggregates. To make that connection we usually have to look at the *system of interaction* between individuals and their environment, that is, between individuals and other individuals or between individuals and the collectivity. And sometimes the results are surprising. Sometimes they are not easily guessed. Sometimes the analysis is difficult. Sometimes it is inconclusive. But even inconclusive analysis can warn against jumping to conclusions about individual intentions from observations of aggregates, or jumping to conclusions about the behavior of aggregates from what one knows or can guess about individual intentions.”⁵

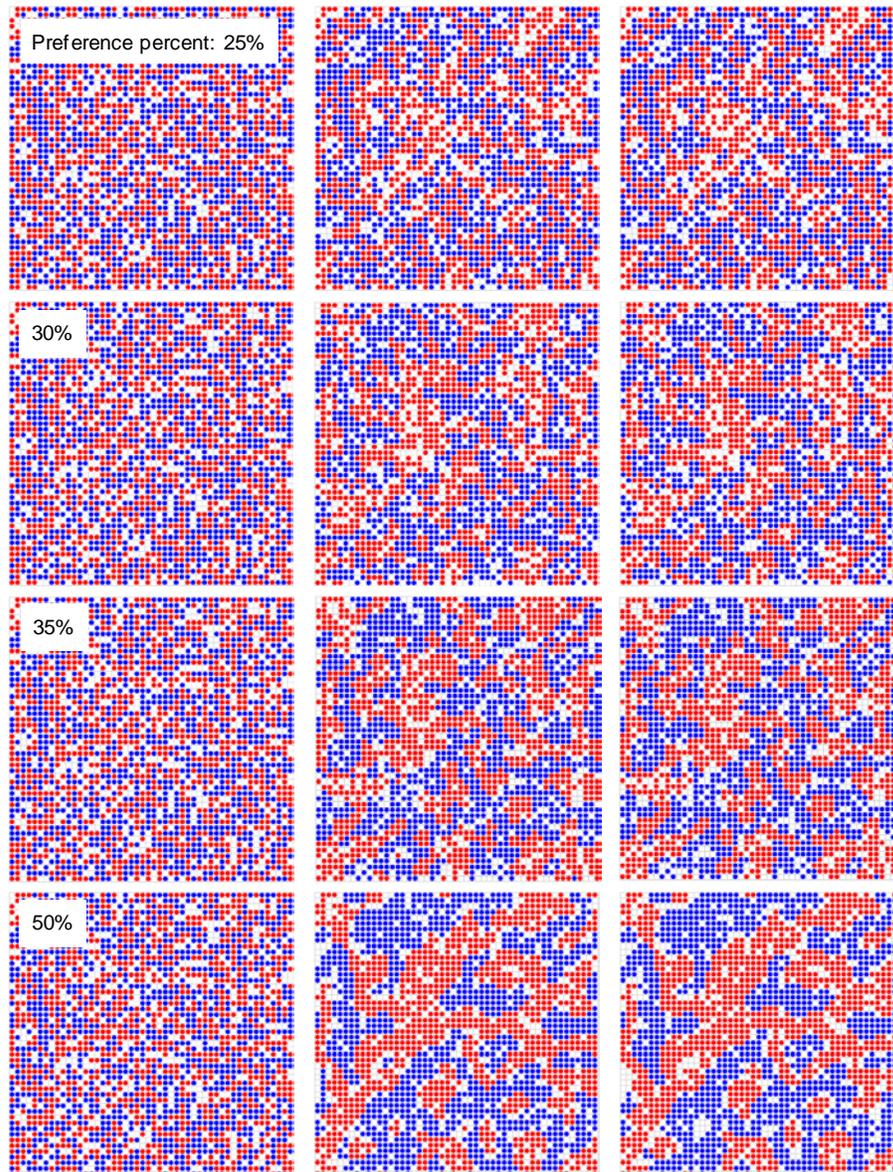
⁵ Schelling (2006), page 14. Note that this was originally written in 1978.

⁶ The classic Schelling segregation model has a slightly different behavior rule: if an agent is dissatisfied, it moves to the *nearest* location where it will be happy.

C. ARTIFICIAL SOCIETY MODELS CONTINUED

1. Schelling segregation model continued

Each row of the following set of Schelling segregation model results shows the agent configuration at time steps 0, 1, and 5 for a particular neighbor preference percentage. For example, the first row shows the results for a neighbor preference percentage of 25 percent.⁷



⁷ Created with the Repast Schelling segregation model accompanying this report, with the following parameters: rows: 50; columns: 50; number of blue agents: 1,000; number of red agents: 1,000; neighbor preference percentages: 0.25, 0.30, 0.35, and 0.50; random number seed: 10.

C. ARTIFICIAL SOCIETY MODELS CONTINUED

1. Schelling segregation model continued

What is remarkable about the model results shown above is that people who want only 50 percent of their neighbors to be like themselves – that is people who are quite tolerant – produce pronounced segregation. Indeed, people who only desire 25 percent like neighbors produce considerable segregation.

The degree of the population’s overall segregation bears little resemblance to the modest biases of its members. Had someone shown you any of the model results above at step 5, would you have guessed the low level of individual bias that gave rise to it?

One might wonder if the model’s counter-intuitive result is simply a product of its lattice environment, or perhaps the homogeneity of neighbor preference percentages among all agents. Interestingly, on other environment topologies, and with asymmetrical distributions of neighbor preference percentages (ie, the preference percentage of blues is different from that of reds), researchers have found that the segregation propensity is even greater. Thus, the result appears to be robust over a wide range of parameters.⁸

The Schelling segregation model is the first model in this report that was developed using the powerful agent-based modeling platform Repast Symphony. To become more comfortable with Repast Symphony, you may enjoy tinkering with the Schelling segregation model (see sidebar). To begin programming agent-based models in Repast Symphony, follow the set-up instruction for Repast Symphony in this report’s accompanying section titled “Getting started with modeling platforms” (located on the report’s SOA web page). You may also find it helpful to read Chapter two: Agent-based modeling.

Exploring the model

To become more comfortable with Repast Symphony models, start by exploring the Schelling segregation model:

- Explore the control buttons at the top of the Repast Symphony environment (hover over the buttons to see a description of what they do). Start by initializing a run, then step through a simulation by clicking on the “step run” button. Notice the successive time step (‘tick’) counts in the upper right-hand corner.
- To move and resize the Schelling segregation environment within its window, right click and drag it into place, then use the wheel on your mouse to expand or contract it (or, alternatively, for a more dramatic effect, hold down the Alt key on your keyboard, hold down the right mouse button, and move the mouse).
- Click on the “toggle info probe” button to hover over any agent and see information about it. Alternatively, double-click on an agent, and its information will appear in a window.
- Click on the “stop run” button to stop the simulation, then click on the “reset” button to reset the simulation. Then go to the ‘parameters’ tab to enter parameters for a new simulation, and click on the “initialize run” button to start another simulation.

⁸ Miller & Page (2007), pages 163-164, and 145-146.

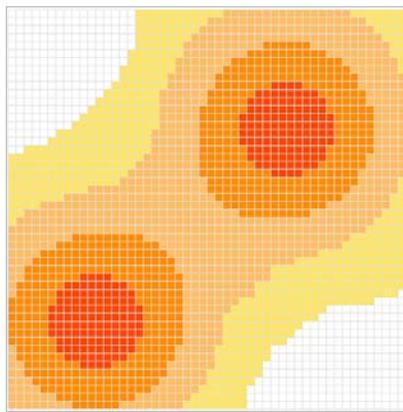
C. ARTIFICIAL SOCIETY MODELS CONTINUED

2. Sugarscape

Now let's explore the artificial society called Sugarscape, where agents can not only move around on an environment, but can also interact with it.

Environment

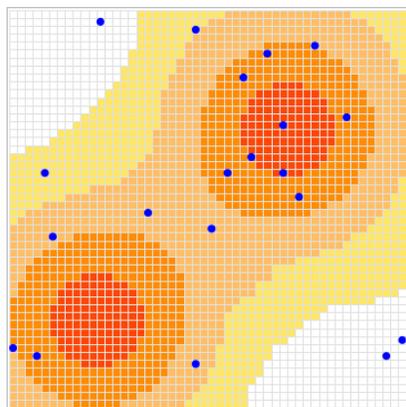
In the simplest version of Sugarscape, the environment consists of a 50 x 50 grid on which there are two mountains of sugar, with higher elevations of sugar represented by darker colors:



The highest elevation is 4 sugar units (the darkest color), followed by terraces of three units, two units, and one unit. Some locations, colored white, have no sugar.

Agents

Into this sweet world are introduced agents at random locations (the blue disks):



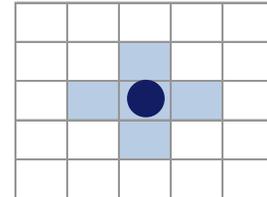
C. ARTIFICIAL SOCIETY MODELS CONTINUED

2. Sugarscape continued

Agent attributes

Each agent has the following attributes:

- **Metabolic rate:** The number of sugar units it must consume each time step in order to stay alive.
- **Vision:** The distance that it can see, measured by a number of grid cells. Agents can only see in north, south, east, and west directions; they cannot see diagonally. For example, the diagram at right highlights the grid cells that an agent with a vision of one can see.
- **Death age:** The age at which the agent dies and is removed from the environment.
- **Location:** Its row and column on the grid (there can be only one agent per location).
- **Wealth:** The amount of sugar units that an agent accumulates.

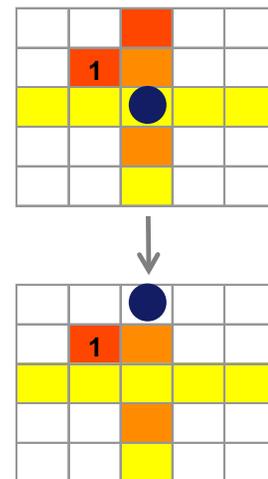


When the simulation starts, each agent is given a random metabolic rate, vision, and death age (all within limits specified by the modeler). These attributes can be thought of as its genetic endowment. At the start, each agent is also given a random location and a random accumulation of sugar units. So, some agents begin at the top of a sugar hill with much wealth, and some begin in the lowlands with no wealth. The agents are thus ‘heterogeneous’; for all their attributes, they can have different values.

Agent behavior rule

At each time step, each agent follows a simple behavior rule: It looks out as far as its vision permits, moves to the nearest unoccupied cell with the most sugar, adds all the sugar units in the cell to its store of sugar wealth, and consumes an amount of sugar equal to its metabolic rate.⁹ If at any time step, the agent does not have enough sugar wealth to satisfy its metabolic rate, it starves to death and is removed from Sugarscape.

For example, consider the agent shown in the upper grid at right. It has a metabolic rate of 1, vision of 2, and – before it moves – wealth of 10. In the next time step, it moves to the location shown (not to location ‘1’, because this location is diagonal), gathers the 4 sugar units, eats 1, and increases its wealth to 13.



⁹ In the behavior rule, there are two random elements: Each time step, agents are selected in random order to follow the behavior rule. Also, if the agent identifies two or more cells to which it could move, it chooses one randomly.

C. ARTIFICIAL SOCIETY MODELS CONTINUED

2. Sugarscape continued

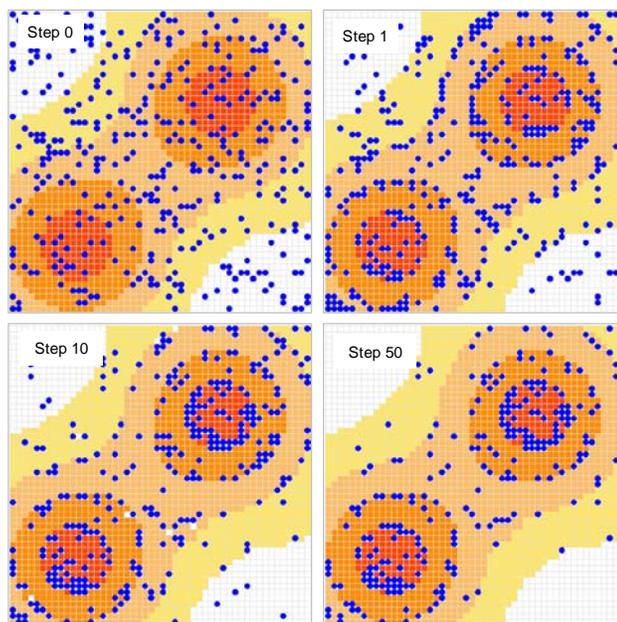
Environment behavior rule

The environment also has a behavior rule: At each time step, the amount of sugar in each cell on the sugar mountain increases by a number of units specified by the modeler (the sugar ‘growth rate’), but the amount in each cell cannot exceed its initial sugar level (4 units at the top, 3 at the next level, etc.).

For example, let the sugar growth rate be 3. If at time step 1 an agent takes all the sugar from a cell at the top of a mountain, at the start of time step 2, the amount of sugar in the cell will grow back to 3.

First Sugarscape simulation

Let’s run our first Sugarscape simulation. If we randomly assign metabolic rates from 1 to 4 inclusive, vision from 1 to 6 inclusive, initial wealth from 5 to 25 inclusive, allow the agents to only die from starvation (not from old age), and set the sugar growth rate to 4 (sugar grows back to its full capacity immediately), can you guess what will happen when we randomly release 400 agents on Sugarscape? Following is the answer, at time steps 0 (initialization), 1, 10, and 50.¹⁰



¹⁰ Created with the Repast Symphony Sugarscape model accompanying this report, with the following parameters: Number of initial agents: 400; maximum vision: 6; maximum metabolic rate: 4; minimum death age: 1000; maximum death age: 1000; minimum initial wealth: 5; maximum initial wealth: 25; sugar growth rate: 4; replace dead agents?: unchecked (no); default random seed: 10.

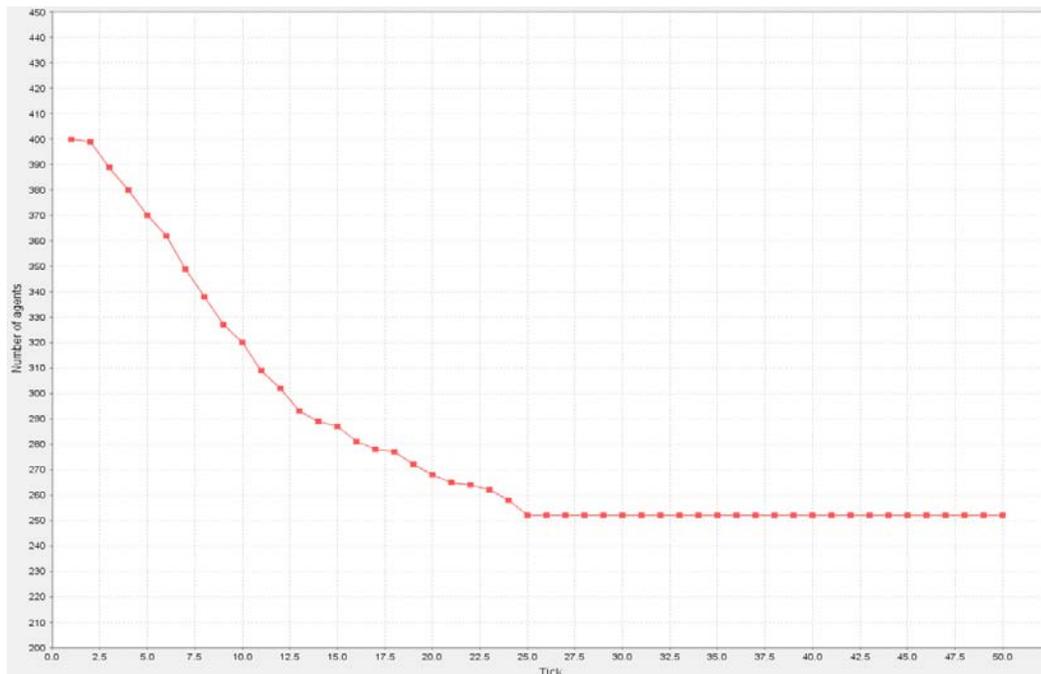
C. ARTIFICIAL SOCIETY MODELS CONTINUED

2. Sugarscape continued

First Sugarscape simulation continued

Two features of the first simulation are immediately striking:

- The agents gravitate to terrace edges, and there they stay. Upon reflection, this behavior makes sense: Because the environment immediately replenishes the sugar supply to full capacity, once agents reach the edge of a terrace, most cannot see any cells with more sugar, and so they don't move.
- Many agents die of starvation. For agents who are born with high metabolism and low vision, and who had the misfortune to be born at lower elevations, life on Sugarscape is hard. In fact, as the chart below shows, the number of agents drops dramatically from 400 to about 250, and then stays at this level. Under the first simulation's assumptions, this is Sugarscape's 'carrying capacity'.¹¹



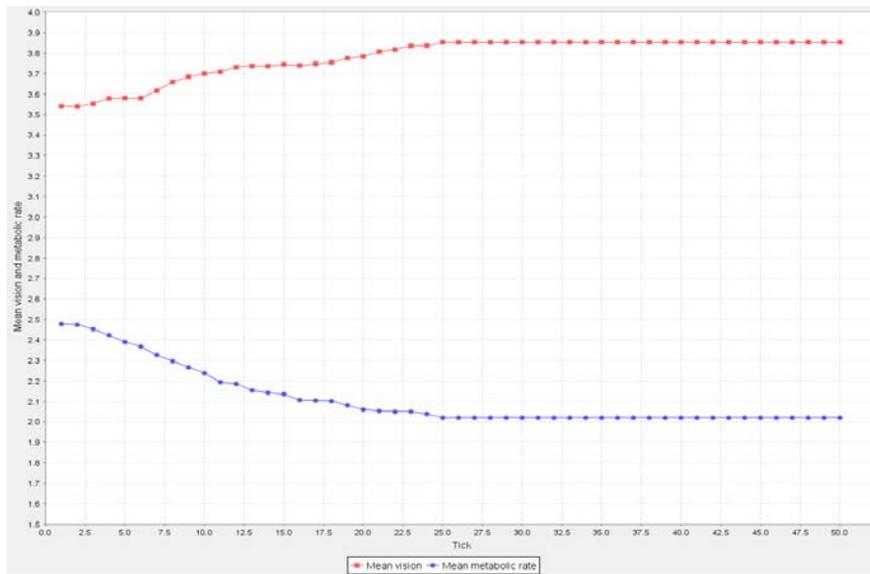
¹¹ This chart was automatically produced by the Repast Symphony Sugarscape model.

C. ARTIFICIAL SOCIETY MODELS CONTINUED

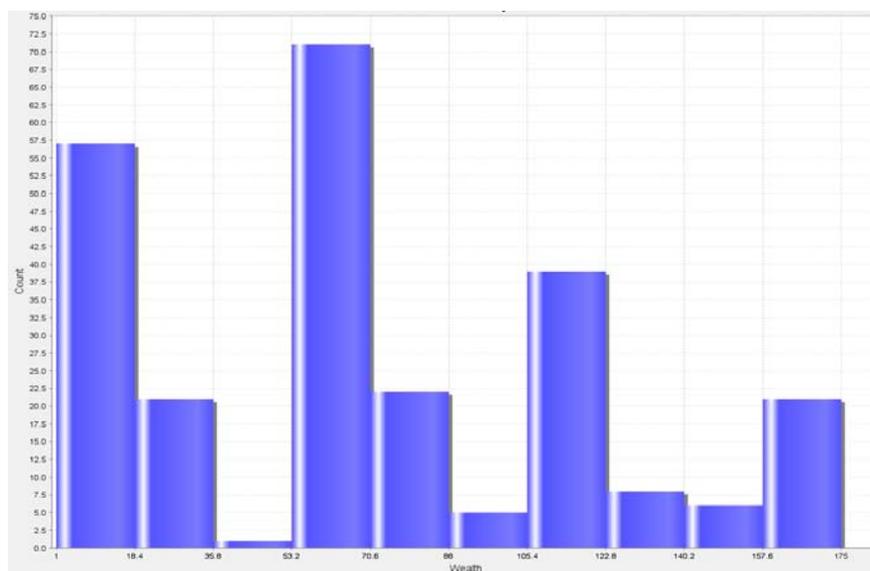
2. Sugarscape continued

First Sugarscape simulation continued

The chart below shows the evolution of the mean vision and metabolic rate over the course of the first simulation. Not surprisingly, evolution favors better vision and a lower metabolic rate.



And the chart below is a histogram of agents' sugar wealth distribution at time step 50. Again, there is nothing particularly surprising.¹²



¹² These charts are automatically produced by the Repast Symphony Sugarscape model.

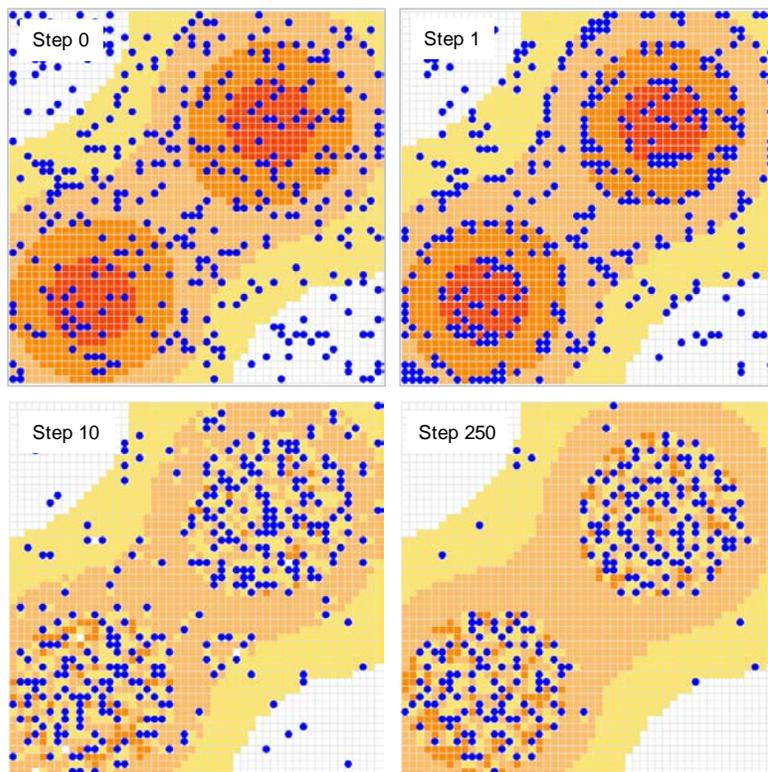
C. ARTIFICIAL SOCIETY MODELS CONTINUED

2. Sugarscape continued

Second Sugarscape simulation

Now let's explore the second Sugarscape simulation. Parameters and behavior rules for the second simulation are the same as those for the first, with the following exception: Rather than immediately growing back to full capacity, the environment's sugar only grows back at the rate of one sugar unit per time step. What impact do you think this change makes?

Following is the second simulation at time steps 0, 1, 10, and 250.¹³



Exploring the model

As with the Schelling segregation model, you will learn a lot, and become more comfortable with Repast Symphony, by exploring the Sugarscape model:

- Double-click on agents to see their attributes and how the attributes change over time. In this way, as a simulation progresses, you can follow the attributes of many agents.
- Change the parameters. For example, explore what happens with just one agent as you vary the parameters.
- Run the simulations for thousands of time steps to see whether simulation results converge or diverge.
- Generate new charts by right-clicking on “Charts” under Environment, and supplying appropriate parameters.
- As we will do in the third Sugarscape simulation, create new output by right clicking on “Outputters” under Environment, and supplying appropriate parameters. Then analyze the output using Excel or R.
- Experiment with parameters on the “Run Options” screen, to see how they can make it easier for you to use the model. For example, you may need to increase the tick delay, in order to slow down the simulation (have you noticed that Repast Symphony models are much faster than Excel models?)
- Move the windows of the display to different areas of the screen to create a configuration that makes it easy for you to view model results.
- Click on the camera or camcorder images to export images or movies. Try importing the images and movies into PowerPoint for a presentation.

¹³ Created with the Repast Symphony Sugarscape model accompanying this report, with the following parameters: Number of initial agents: 400; maximum vision: 6; maximum metabolic rate: 4; minimum death age: 1000; maximum death age: 1000; minimum initial wealth: 5; maximum initial wealth: 25; sugar growth rate: 1; replace dead agents?: unchecked (no); default random seed: 10.

C. ARTIFICIAL SOCIETY MODELS CONTINUED

2. Sugarscape continued

Second Sugarscape simulation continued

In this simulation, the agents evolve into two colonies, one atop each mountain. What is perhaps most intriguing: by step 10, the elevation of Sugarscape has been leveled; there are practically no more cells with 4 or even 3 sugar units. Without centralized direction, the agents have evolved into a surprisingly efficient harvesting machine.

As in the first simulation, here the vision increased from about 3.5 to about 3.9, the metabolic rate dropped from about 2.5 to about 1.8, and the number of agents dropped from 400 to 235.

In this overview of Sugarscape, we have only touched its surface. Epstein and Axtell's work went much further, addressing social networks, pollution, effects of inheritance, genealogical networks, cultural dynamics, war, trade, competition, and disease transmission, all of which can be modeled with *Repast Symphony*.

In the next section, we'll explore two practical applications of Sugarscape relating to wealth distribution and non-equilibrium economics.

The point

In the conclusion of *Growing artificial societies*, Epstein and Axtell write:

“The main point of the preceding chapters is simply this: A wide range of important social, or collective, phenomena can be made to emerge from the spatio-temporal interaction of autonomous agents operating on landscapes under simple local rules. ...

As we have demonstrated, in an agent-based model each individual can have a variety of behavioral rules, and these can all be active simultaneously. When such multifaceted agents are released into an environment in which (and with which) they interact, the resulting society will – unavoidably – couple demography, economics, cultural change, conflict, and public health. All these spheres of social life will emerge – and merge – naturally and without top-down specification, from the purely local interactions of the individual agents. Because the individual is multifaceted, so is the society. The fixed coefficients of aggregate models – such as fertility rates or savings rates – become dynamic, emergent entities in bottom-up models.”¹⁴

¹⁴ Epstein & Axtell (1996), pages 153 and 158.

D. PRACTICAL APPLICATIONS

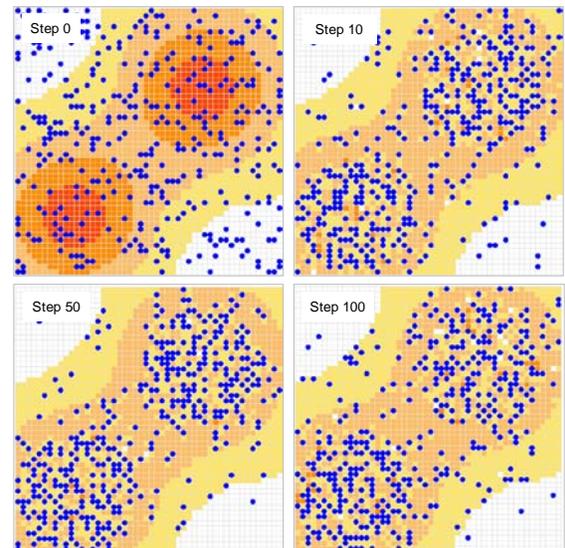
As with networks and cellular automata, my search of actuarial databases for applications of artificial societies came up empty. In fact, it appears that applications of artificial societies that are relevant to the actuarial domain are relatively rare. Following are three such applications, one from Sugarscape regarding wealth distribution, another from Sugarscape regarding non-equilibrium economics, one about epidemic dynamics that may interest health actuaries, and one called *Archimedes* about healthcare decision making that is sure to interest both health and non-health actuaries.

1. Wealth distribution

In 1895, Vilfredo Pareto collected income data from a number of countries, and found that their frequency distribution followed an unusual pattern, what we now call the Pareto distribution, or power-law distribution. Rather than the normal distribution of income that one might expect, income is highly skewed: many people have relatively low income, while the income of a very few is vast. Since then, traditional economists have confirmed that both income and wealth follow a power-law distribution. But, despite considerable work, they could never figure out *why*. The simple agents of Sugarscape provide an explanation.

For this application of Sugarscape, we introduce death and birth: We set the minimum and maximum ages of natural death at 60 and 100, and when an agent dies from natural death or starvation, we replace it with a newborn agent who has a random initial genetic endowment (vision and metabolic rate), a random death age (between 60 and 100), and a random initial location. Otherwise, the parameters are the same as for the second simulation. At right are results at time steps 1, 10, 50, and 100.¹⁵

There is nothing obviously new in these visual results. As well, analysis of agent counts and attributes are not surprising: the number of agents stays steady at 250 (because we replace each dead agent with a newborn), and the genetic attributes of vision and metabolic rates again evolve to favor survival.

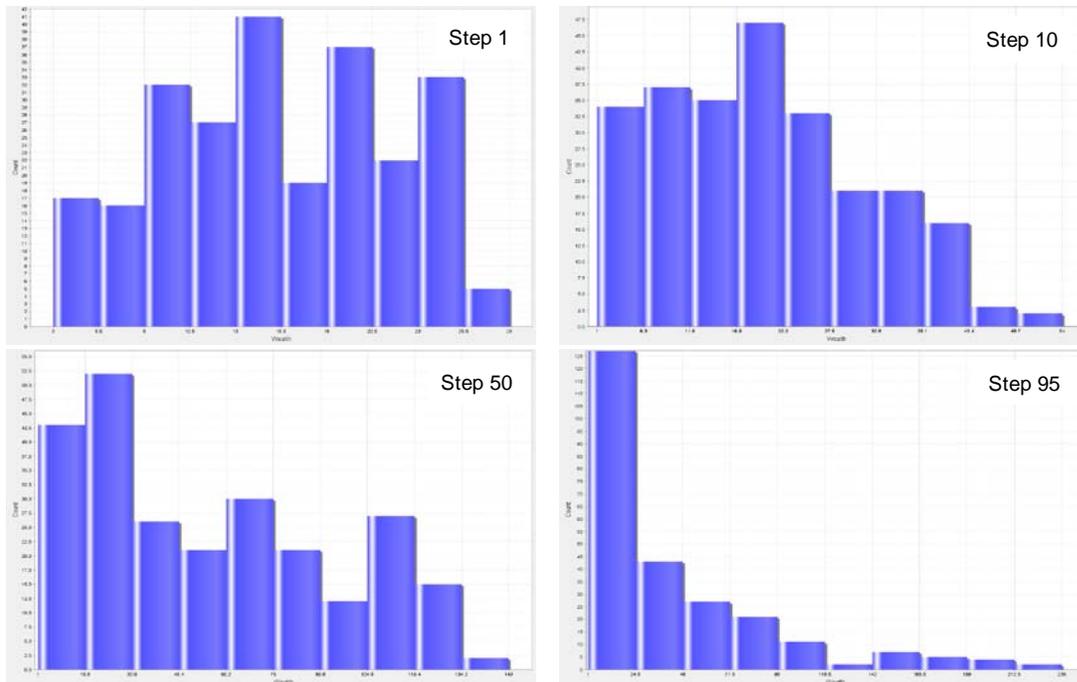


¹⁵ Created with the Repast Symphony Sugarscape model accompanying this report, with the following parameters: Number of initial agents: 250 (the carrying capacity of Sugarscape); maximum vision: 6; maximum metabolic rate: 4; minimum death age: 60; maximum death age: 100; minimum initial wealth: 5; maximum initial wealth: 25; sugar growth rate: 1; replace dead agents?: checked (yes); default random seed: 10.

D. PRACTICAL APPLICATIONS CONTINUED

1. Wealth distribution continued

But there is a new phenomenon – a result that blasted traditional economists awake – illustrated by the following charts of agent wealth distribution at time steps 1, 10, 50, and 95:¹⁶



Now that you know all the rules of Sugarscape, how would you explain why its inhabitants end up with a skewed distribution of wealth? How would you explain why two agents – both with average vision, metabolic rate, death age, and initial location – end up on opposite ends of the wealth spectrum? How would you help to develop a policy to reduce such wealth inequality?

Would you apply a traditional method such as regression analysis, and determine the combination of factors that correlate most strongly with wealth, in an attempt to discover cause-effect relationships? Would you data-mine the results to find an answer, or apply predictive analytics?

¹⁶ These charts are automatically produced by the Repast Symphony Sugarscape model.

D. PRACTICAL APPLICATIONS CONTINUED

1. Wealth distribution continued

Unfortunately, such traditional methods would fail. Because agent attributes (genetic endowments, initial wealth, and location) are initially distributed randomly and uniformly, if the wealth distribution were dependent on attributes alone, one would logically find wealth roughly evenly distributed by attribute. Thus, no attribute mining or analysis will discover a linear cause-effect relationship. Nor are simplistic political views such as “the rich exploit the poor to get richer” or “it’s the stupid lazy people who are poor” accurate or useful.

The true explanation is deeper. The cause of the skewed wealth distribution is the system as a whole; it is *everything*. It is Sugarscape’s structure, the agent endowments, and the behavior rules all together; and, perhaps most interestingly, it is luck. The skewed wealth distribution is simply an emergent property of the system (see sidebar), and cannot be reduced to simple cause-effect explanations.

The reason for horizontal inequality – why two agents with similar initial endowments end up on opposite ends of the wealth distribution – is simply luck: one agent happened to turn south toward an area of low sugar, whereas the other turned north to riches.

This understanding has important implications for policy formation. Simulation models such as Sugarscape would enable regulators and other policy makers to understand which changes in environmental structure or behavior rules can have the most potent effects. In fact, in the area of energy regulation such models are beginning to enter the regulatory process: To avoid repeating the 2000 disaster when Enron and other companies manipulated energy supplies and prices, several US states now use agent-based models to test complex electricity market designs before implementation.¹⁸ The point of these models is not prediction, but understanding.

Emergence of wealth

Epstein and Axtell write about the distribution of wealth on Sugarscape:

“In the sciences of complexity, we would call this skewed distribution an *emergent structure*, a stable macroscopic or aggregate pattern induced by the local interaction of the agents. Since it emerged ‘from the bottom up,’ we point to it as an example of *self-organization*. Left to their own, strictly local, devices the agents achieve a *collective* structuring of some sort. This distribution is our first example of a so-called *emergent structure*.”

The term ‘emergence’ appears in certain areas of complexity theory, distributed artificial intelligence, and philosophy. It is used in a variety of ways to describe situations in which the interaction of many autonomous individual components produces some kind of coherent, systematic behavior involving multiple agents. To our knowledge, no completely satisfactory formal theory of ‘emergence’ has been given. A particularly loose usage of ‘emergent’ simply equates it with ‘surprising’ or ‘unexpected’, as when researchers are unprepared for the kind of systematic behavior that emanates from their computers. A less subjective usage applies the term to group behaviors that are qualitatively different from the behaviors of individuals composing the group.

We use the term ‘emergent’ to denote *stable macroscopic patterns arising from the local interaction of agents*. One example is the skewed wealth distribution; here, emergent structure is statistical in nature.”¹⁷

¹⁷ Epstein & Axtell (1996), pages 34-35.

¹⁸ Economist Leigh Teftatson of Iowa University (in Ames, Iowa) has led the development of the agent-based model known as the *Ames wholesale power market test bed*.

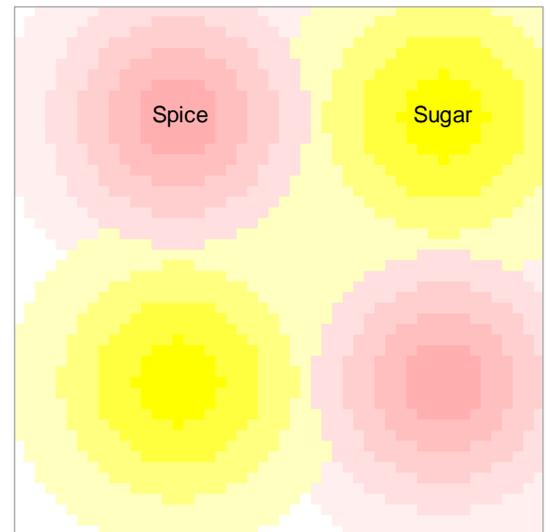
D. PRACTICAL APPLICATIONS CONTINUED

2. Non-equilibrium economics

Traditional economics is elegant theory, but in key areas it doesn't square with reality. For example, contrary to what traditional economic theory predicts, in the real world there is horizontal inequality, multiple prices in a market (rather than a 'law of one price'), and relatively high levels of price volatility and trading volume (rather than equilibrium). To the amazement – and consternation – of traditional economists, not only does Sugarscape reproduce fundamental economic results such as the law of supply and demand and welfare gains from trade, it also explains key real-world phenomena that traditional economics cannot explain, all with simple boundedly-rational agents following simple behavior rules. Let's see how Epstein and Axtell accomplished this.

Environment

To enable trade on Sugarscape, a second resource – spice – is added. In addition to two mountains of sugar, two similar mountains of spice are added, but on the SE-NW diagonal (see the diagram to the right). With this arrangement, most cells of Sugarscape now have units of both sugar and spice.



Agents

To the agent attributes are added:

- **Spice wealth:** the agent's accumulation of spice.
- **Spice metabolism:** the units of spice that an agent must consume each time step in order to stay alive. Agents die if their sugar or spice accumulation falls to zero.

Agent behavior

Agents attempt to gather the resources (sugar or spice) they need to stay alive. As in the one-resource case, an agent looks around to the extent of its vision for a cell that maximizes its welfare. Specifically, an agent chooses the cell that maximizes the welfare quantity:

$$(w_1 + x_1^c)^{\frac{m_1}{m_1+m_2}} \times (w_2 + x_2^c)^{\frac{m_2}{m_1+m_2}}$$

where x_1^c and x_2^c are the sugar and spice levels of cell c , w_1 and w_2 are the agent's wealth levels of sugar and spice, and m_1 and m_2 are the agent's metabolism rates for sugar and spice.¹⁹

¹⁹ The quantity is a Cobb-Douglas functional form, widely used in economic theory.

D. PRACTICAL APPLICATIONS CONTINUED

2. Non-equilibrium economics continued

Agent behavior continued

To make sense of this welfare quantity, assume that the agent's metabolic rate for sugar is twice that for spice ($m_1 = 2m_2$), and that its wealth levels of sugar and spice are equal ($w_1 = w_2$). Then, if the agent has a choice between two cells with the same level of spice, because the quantity would be larger for the cell with the higher sugar units, the agent would choose the cell with more sugar.

As in the case of a single resource, if there are two cells that produce the same welfare quantity, the agent chooses the nearest cell.

In addition, agents can now trade. At each time step, each agent randomly chooses one of its von Neumann neighbors as a trade partner. When two agents trade, each agent first computes the value of its sugar and spice stores as:

$$V = \frac{w_2/m_2}{w_1/m_1}$$

which is the ratio of the time steps to death if the agent gathers no more spice to the time steps to death if it gathers no more sugar.

If for two trading agents A and B this value is equal ($V_A = V_B$) they don't trade. If $V_A > V_B$, agent A buys sugar and sells spice.

The trading price is

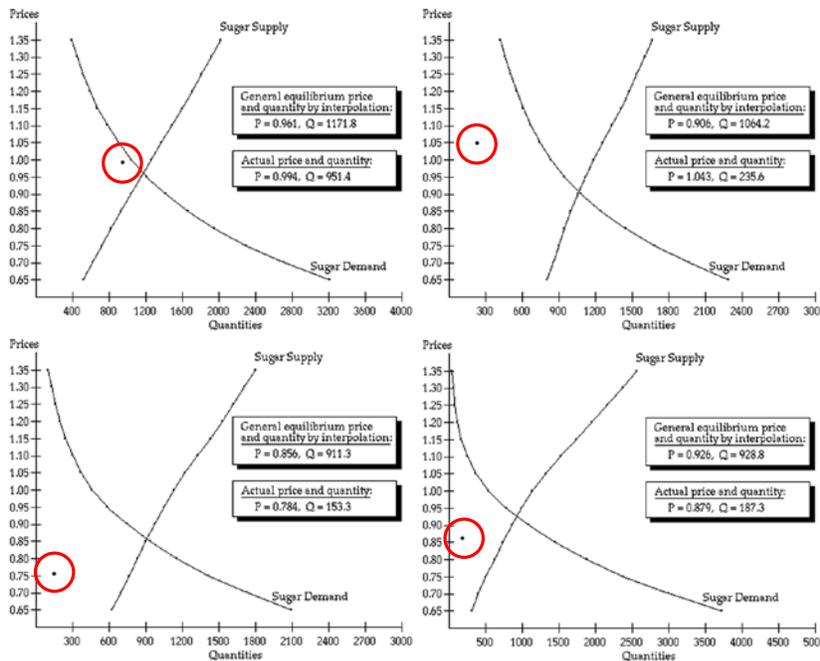
$$p = \sqrt{V_A \times V_B}$$

If $p > 1$, p units of spice are exchanged for 1 unit of sugar. If $p < 1$, then one unit of spice is exchanged for $1/p$ units of sugar. If such an exchange increases the welfare of both agents but does not change the initial relative relationship between V_A and V_B , they go ahead and trade. The agents then recalculate their V and continue trading until further exchange does not increase their mutual welfare.

D. PRACTICAL APPLICATIONS CONTINUED

2. Non-equilibrium economics continued

Now, let's see the impact of the new agent attributes, behavior rules, and environment. The following charts show supply and demand curves from Sugarscape, as well as the actual price and quantity at four times during a trade simulation.²¹



Contrary to traditional economic theory, the price and quantity traded do not correspond to the intersection of supply and demand curves. As the charts show, while the actual price moves around the general equilibrium price, the actual quantity traded is always less than what is necessary to 'clear the market' and attain optimum welfare. There is always some demand unfulfilled.

The model demonstrates that the reason for mismatch between Sugarscape's actual price and quantity, and the intersection of supply and demand curves, is timing. At each time step, agents produce (gather sugar and spice), consume, and trade. But, it takes time for the system as a whole to reach price equilibrium. By the time the society as a whole converges to price equilibrium, agents are in new positions of production and consumption, shifting the supply/demand curves intersection to a new place. The actual price and quantity forever trails the intersection point.

²⁰ Buchanan (2009)

²¹ From a video clip supplied with Epstein & Axtell (1996), based on 200 initial agents, initial endowments of sugar and spice between 25 and 50, vision between 1 and 5, and metabolism rates for sugar and spice between 1 and 5.

Without presupposing the result

"Financial regulators do not have the tools they need to predict and prevent meltdowns ... They can do a good job of tracking an economy using the statistical measures of standard economics, as long as the influences on the economy are independent of each other, and the past remains a reliable guide to the future. But the recent financial collapse was a 'systemic' meltdown, in which intertwined breakdowns in housing, banking and many other sectors conspired to destabilize the system as a whole. And the past has been anything but a reliable guide of late: witness how US analysts were led astray by decades of data suggesting that housing values would never simultaneously fall across the nation.

Likewise, economists can get reasonably good insights by assuming that human behavior leads to stable, self-regulating markets, with the prices of stocks, houses and other things never departing too far from equilibrium. But 'stability' is a word few would use to describe the chaotic markets of the past few years, when complex, nonlinear feedbacks fuelled the boom and bust of the dot-com and housing bubbles, and when banks took extreme risks in pursuit of ever higher profits.

In an effort to deal with such messy realities, a few economists – often working with physicists and others outside the economic mainstream – have spent the past decade or so exploring 'agent-based' models that make only minimal assumptions about human behavior or inherent market stability. The idea is to build a virtual market in a computer and populate it with artificially intelligent bits of software – 'agents' – that interact with one another much as people do in a real market. The computer then lets the overall behavior of the market emerge from the actions of the individual agents, without presupposing the result."

Mark Buchanan²⁰

D. PRACTICAL APPLICATIONS CONTINUED

3. Epidemic dynamics

In the U.S., concern about bioterrorism has centered on smallpox. Smallpox is highly communicable and kills about 30 percent of those infected. With the current low level of immunization against this disease in the U.S. (smallpox vaccinations ended in 1972) it poses a significant threat.

In the wake of the September 11 attacks, there was considerable debate about an appropriate national strategy to counter smallpox bioterror. In 2003 and 2004, Epstein et al developed a virtual-world model to help define such a strategy.²²

Agents

The model consists of 800 agents, representing individuals who live and work in two towns, Circletown and Squaretown. The agents of Circletown are represented by circular disks, and those of Squaretown by squares.

Each agent has the following disease states (represented by the following colors):

-  Healthy and susceptible
-  Infected, asymptomatic, noncontagious
-  Infected, mild symptoms, slightly contagious
-  Infected, severe symptoms, highly contagious
-  Recovered, immune
-  Dead

Each also has the following attributes: city ID, family ID, daytime role (worker, student, hospital worker), and workplace or school ID (depending on whether the agent is an adult or a child).

²² *Toward a containment strategy for smallpox bioterror: an individual-based computational approach*, Chapter 12 of Epstein (2006).

Complexity science – an introduction (and invitation) for actuaries

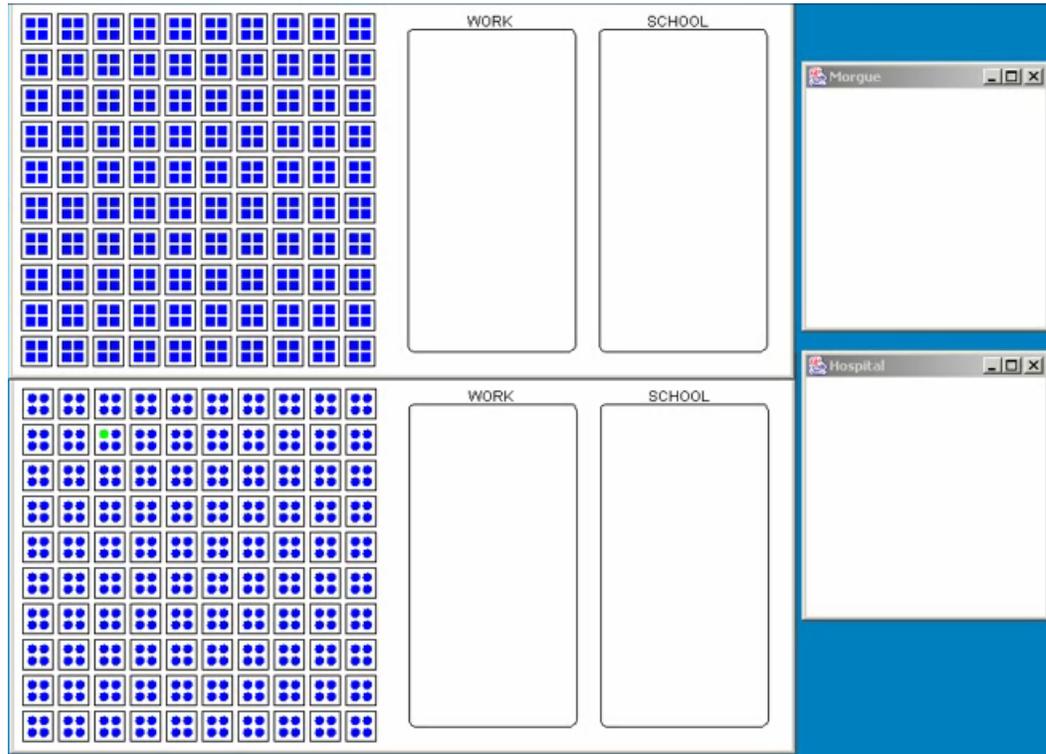
Chapter six: Artificial societies continued

E. PRACTICAL APPLICATIONS CONTINUED

3. Epidemic dynamics continued

Environment

The model's environment, with its 800 agents is shown below.



On the left are the two towns, Squaretown on top, and Circletown on bottom. It is currently night time, and all agents are at home with their families (each of which consists of two parents and two children). In the center are the workplaces and schools for the two towns, one of each per town. And on the right is the hospital and the morgue, serving both towns.

Time

Each time step represents one hour. There are 20 hours in each day, 10 of which are day time (when inhabitants are in school or at work if they are healthy), and 10 of which are night time (when inhabitants are at home).

E. PRACTICAL APPLICATIONS CONTINUED

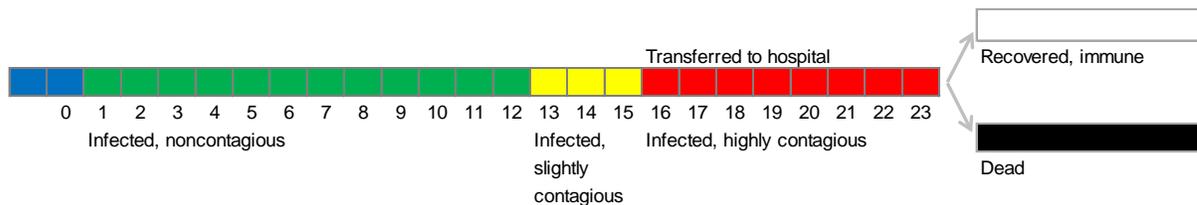
3. Epidemic dynamics continued

Agent behavior rules

At the start of each day, all healthy agents go either to work or school. Each day, an agent’s location at work or school is randomly determined (and so its work and school neighbors vary from day to day). Ten percent of adults from each town commute to work in the other town, but all children go to school in their town. Five agents from each town work in the hospital. At night, all agents return home.

At each time step, each agent (in a randomly-determined order) has an interaction with a randomly-selected neighbor in its Moore neighborhood (where at night the Moore neighborhood of an agent is understood to consist only of its family). Some of these interactions will result in a ‘contact’ (at work and school, 30 percent of interactions result in a contact; in the hospital, 100 percent do.) If an agent has contact with an agent in the slightly contagious phase of smallpox, the agent will contract the disease with a probability of 0.20. During the highly contagious phase, the probability is 0.40.

For an infected agent, the slightly contagious phase lasts for days 13 through 15 after contracting the disease, and the highly contagious phase lasts for days 16 through 23. At day 16, infected agents are hospitalized, and from then through day 23, they can die with cumulative probability of 0.30. Dead agents are placed in the morgue.



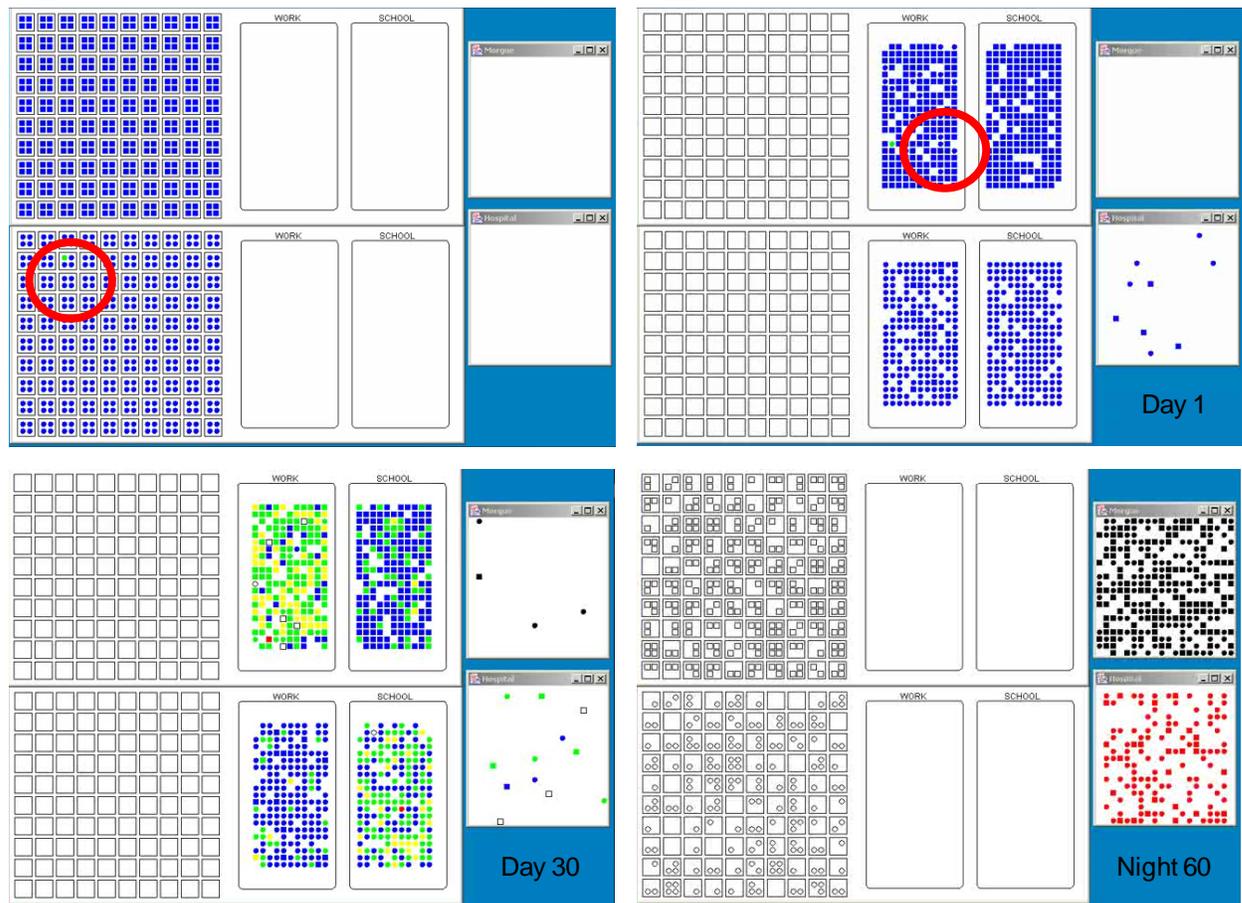
The model’s parameters for disease transmission were obtained by calibrating model results to historical data from 49 instances of a single smallpox case being introduced into Europe during the period 1950-1971. To calibrate the data, the researchers ran about 10,000 simulations, sweeping through all the major model parameter combinations. For the model, they chose parameters that minimized the sum of squared deviations from historical data.

E. PRACTICAL APPLICATIONS CONTINUED

3. Epidemic dynamics continued

Simulation

Starting with one infected agent, and without any vaccination strategy, the following sequence shows the simulation at times 0 (day 10 of the initial agent's infection), day 1, day 30, and night 60. As you see, by night 60, the towns have been decimated by disease.



By introducing a variety of immunization strategies into such simulations, Epstein discovered one that was superior to those previously considered, namely:

- Vaccination of 100 percent of hospital workers
- Voluntary revaccination of healthy individuals successfully vaccinated in the past.
- Hospital isolation of confirmed cases.
- Vaccination of household members of confirmed cases.

Again, the point is to understand, not to predict.

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions

The closest thing we have to a model of the US healthcare system is Archimedes. Twenty years and tens of millions in the making, Archimedes is an agent-based model that aids healthcare decision makers of every kind. Conceived and nurtured by David Eddy (see sidebar) and Leonard Schlessinger (PhD in physics), it gathers together myriad strands of evidence about human physiology, diseases, diagnoses, treatments, physician behavior, patient behavior, and healthcare organizational logistics, and weaves them into a quantitative tapestry to help decision makers see likely outcomes of their choices.

This section describes the model and gives one example of its use. Because the model's scope is wide, and its architecture deep, this description is longer than others in the report. Even if you are not a health actuary, you may find it enlightening, for Archimedes is an excellent example of how a complex social system model is conceived, developed, validated, and used.

Scope

Archimedes was designed to support a wide spectrum of US healthcare decisions, including:

- **Selecting new treatment combinations.** Archimedes helps decision makers understand the potential impact of new treatment combinations. It can model the potential outcome of two treatments that have been researched separately but never implemented in combination. For example, in 2003 Archimedes showed that a simple combination of aspirin and generic drugs to lower blood pressure and cholesterol would dramatically reduce the heart attacks and strokes that commonly accompany diabetes. David Eddy convinced the Kaiser Permanente health plan to change the treatment of diabetics to incorporate the new treatment combination, and thereby improved patient outcomes and reduced costs by hundreds of millions, just as Archimedes projected.²⁴
- **Focusing data collection.** Today much healthcare data is collected through research and clinical trials, but it is largely a hodge-podge. By organizing healthcare data in one model, Archimedes highlights important data gaps and focuses data collection efforts.

²³ Carey (2006)

²⁴ Carey (2006)

David and the US healthcare system

David Eddy, an MD trained in heart surgery with a PhD in applied mathematics, is an iconoclast. He has devoted his career to exposing the darkest secret of the US healthcare system. About this secret, he says, “The problem is that we don’t know what we are doing. ... I’ve spent about 25 years proving that what we lovingly call clinical judgment is woefully outmatched by the complexities of medicine.”²³ His slingshot is Archimedes.

The reason why so much of US health care is ineffective, Dr. Eddy claims, is that healthcare providers and decision makers cannot make sense of the vast complexities of human biology, diagnostic procedures, treatments, and cascading new medical knowledge in order to treat patients effectively and consistently. Consequently, they often rely on overly-simplistic heuristics to make medical decisions.

His PhD dissertation overturned two guidelines of the time. He showed that:

- Annual chest X-rays are worthless.
- Yearly Pap smears for women at low risk of cervical cancer are a waste of resources.

He also showed that bone marrow transplants for breast cancer don’t work, and traced one common practice – preventing women from giving birth vaginally if they had previously had a cesarean – to the erroneous recommendation of one lone doctor.

He coined the term ‘evidence-based medicine’ in the 1980s, and has spent the last 20 years developing Archimedes to help healthcare decision makers make decisions based on evidence. He is deservedly proud of the model, and hopes that someday every decision maker will pause before making a healthcare decision and ask, “What does Archimedes say?”

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Scope continued

- **Comparing effectiveness.** Archimedes can compare the costs and effectiveness of various healthcare interventions such as diagnostic procedures, treatments, and preventive procedures. Such comparisons are useful to decision makers for developing guidelines, setting priorities and designing performance measures.
- **Developing incentives.** The model can help decision makers develop incentives to improve provider performance and patient compliance.
- **Designing health plans.** The model can help decision makers understand the potential cost and outcome impacts of various health plan designs.
- **Developing new care processes.** The model can determine the impact on costs and outcomes of changing guidelines, logistics, and timing of healthcare processes. Such results are important to making decisions for disease management and quality improvement programs.
- **Extending research results.** Archimedes can extend research results to populations that are different from the original research population, such as to people with more severe diseases or with different risk factors. It can also project results of short-term research into the future, and the results of a program observed in one setting to other settings.

Archimedes supports decisions where controlled trials do not exist. Many decisions cannot be studied through controlled trials, because such trials would be too expensive, too time-consuming, or impossible. For example, the Archimedes results for diabetes screening discussed at the end of this section could not have been obtained through a controlled trial. Dr. Eddy emphasizes that no one has ever intended for Archimedes to replace controlled trials.

Archimedes models the following diseases and conditions:

- | | | |
|----------------------------|---------------------------------|----------------------|
| ▪ asthma | ▪ coronary artery disease (CAD) | ▪ lung cancer |
| ▪ breast cancer | ▪ diabetes (types I and II) | ▪ metabolic syndrome |
| ▪ colon cancer | ▪ dyslipidemia | ▪ obesity |
| ▪ congestive heart failure | ▪ hypertension | ▪ stroke |

The diseases it first modeled (and still its signature) are CAD and diabetes. Adding a new disease to the model takes six months to one year.

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Agents

For the diseases that Archimedes models, it includes all the types of agents and agent attributes that healthcare decision makers consider important: all variables that physicians use in caring for patients, and all agents and variables that healthcare organizations consider in determining cost and efficacy. That’s a lot of agents (dozens in fact) and attributes (hundreds).²⁵

For example, Archimedes includes many types of providers, such as nurses, radiologists, and physicians; and within each type there are sub-types and sometimes sub-sub-types, such as surgeons and cardiac surgeons. It also includes operations personnel such as healthcare hotline operators.

For provider agents it includes attributes such as salary, schedule, and skill level. For patients, it includes hundreds of attributes such as:

<p>Demographics age gender race/ethnicity location education occupation</p>	<p>Body measures height weight waist circumference hip circumference body mass index artery occlusion %</p>	<p>Body chemistry LDL HDL total cholesterol triglycerides blood glucose HbA1c</p>	<p>Risk factors smoking family history genetic profile alcohol exercise stress</p>
<p>Signs/symptoms fatigue chest pain thirst blurred vision</p>	<p>Metabolism insulin resistance glucose production glucose tolerance insulin production</p>	<p>Vital signs blood pressure heart rate temperature respiratory rate</p>	<p>Events visits admissions contacts health outcome</p>
<p>Procedures chest x-ray mammogram heart surgery lab test</p>	<p>Diagnoses type 1 diabetes asthma cancer stroke</p>	<p>Treatments anti-hypertensives nitroglycerin aspirin diuretics</p>	<p>Lifestyle exercise diet sleep drug rehab</p>

The value of an agent’s attributes can be determined at any time during an agent’s life. For example, Archimedes tracks the percent occlusion of an agent’s coronary arteries throughout the agent’s life. To accomplish this, all time-dependent attributes are represented as continuous functions of time.

²⁵ Much of the description of Archimedes is obtained from an excellent webcast presented by Dr. Eddy. To see the webcast, go to www.archimedesmodel.com/webinar and click on ‘download archive’.

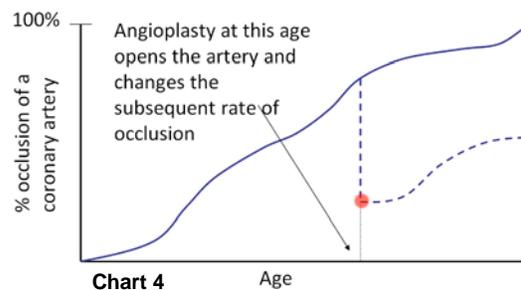
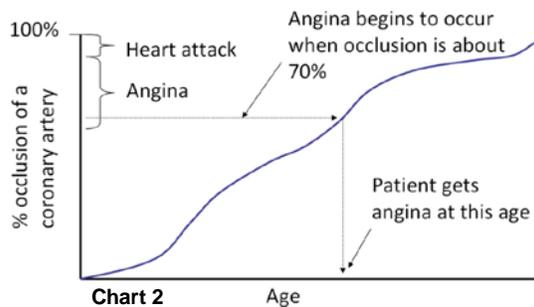
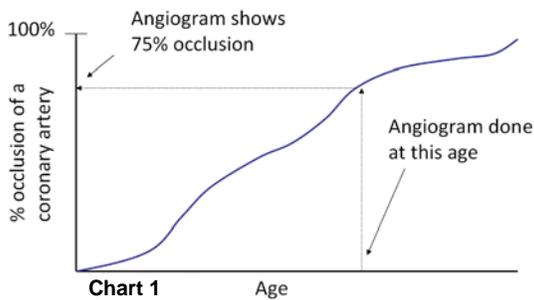
D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Agents continued

For example, the series of charts below illustrates how the agent attribute ‘percent occlusion of a coronary artery’ is developed and used. The percent occlusion is represented by a continuous function of time (age) from age 0 to any relevant age until death. The function in the charts below is represented by the blue curve.

For example, an attribute function might be represented by the polynomial $a + bt^2 + ct^3 + dt^4 + et^5$, where $a - e$ are constants and t represents time.²⁶ The constants are generally derived from population research studies. For example, the percent occlusion function might come from US population studies of arterial occlusion, broken down by age, gender, race, and other population characteristics.



Just as the histories of arterial occlusion are unique for people in the real world, the attribute function for each instantiation of an agent in an Archimedes simulation is unique, based on the agent’s other attributes and random perturbation. For instance, one person’s percent occlusion might remain below 50 percent, while another person’s might reach 100 percent at an early age.

²⁶ Schlessinger & Eddy (2001), page 40.

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Agents continued

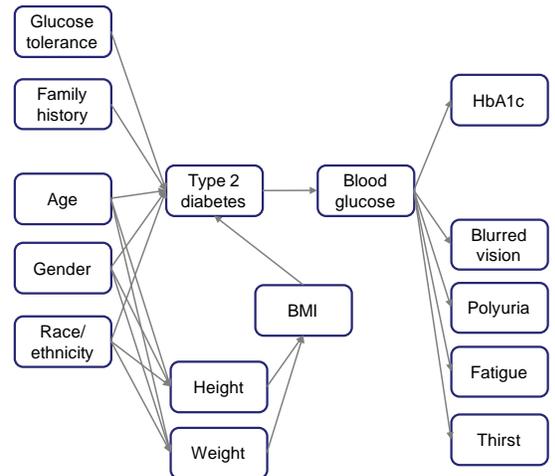
As Chart 1 on the previous page illustrates, if an agent’s percent occlusion is probed at any time, such as with an angiogram test, Archimedes returns the value of the agent’s attribute function at that time. As an interesting side note: To make Archimedes as realistic as possible, its developers introduced real-world error into test results. So, the angiogram might result in a value for the attribute that is different than its actual value.

As Chart 2 shows, an agent’s symptoms can be a function of other attributes. For example, angina (chest pain) for a certain agent could be defined to start when the percent occlusion reaches 70 percent. Another agent might not experience angina until occlusion is 90 percent.

Charts 3 and 4 show how an attribute function can be altered by treatment. Cholesterol lowering treatment reduces the rate of occlusion progression, while surgery produces an immediate drop in the occlusion percent.

Because patient attributes are such a vital component of Archimedes, we will explore them more deeply. For this purpose, let’s look at attributes related to type 2 diabetes. (Type 2 diabetes – also called ‘adult-onset diabetes’ – is a major worldwide health problem in which the body fails to properly use insulin to control blood glucose, causing it to rise to health-threatening levels.)

For the patient attributes related to type 2 diabetes, the diagram at right shows their interrelationships.²⁷ The boxes are the attributes, and the lines are equations (generally continuous equations as a function of time) relating the attributes. For example, the attribute function for the diagnoses of type 2 diabetes depends on the demographic attributes age, gender, and race/ethnicity, as well as on the agent’s family history of diabetes, its glucose tolerance, and its body mass index (BMI), which in turn depends on height and weight. Similarly, the agent’s symptoms (fatigue, thirst, etc.) and body chemistry (blood glucose and HbA1c – hemoglobin with glucose) depend on the type 2 diabetes attribute.²⁸



²⁷ The diagram is simplified. For a more complete diagram, see J. Kahn (2009).

²⁸ Eddy & Schlessinger (2003)

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Agents continued

As an example of a continuous attribute function in Archimedes, following is the type 2 diabetes attribute function:²⁹

$$\left(1 - \exp\left(\frac{-a \times IGT}{1 + \exp\left(-\frac{t-b}{c}\right)}\right)\right) \times CBMI(t)/\varepsilon$$

where:

- t is time (age in days)
- IGT is a random number between 0 and 2 representing the agent’s risk of impaired glucose tolerance
- CBMI(t) is the contribution to diabetes progression due to BMI, as a function of time – another continuous function
- ε is a random number between 0 and 1, to reflect the randomness of diabetes development in real populations
- a, b, and c are constants

The constants are derived from population studies of type 2 diabetes that produce tables such as that shown at right, which is based on the NHANES (National Health and Nutrition Examination Survey).³⁰

Race/sex	Age (years)				
	20-39	40-49	50-59	60-74	≥75
All races†					
Both sexes	1.1	3.9	8.0	12.6	13.2
Men	1.1	3.3	9.6	11.8	13.8
Women	1.1	4.4	6.6	13.3	12.8
Non-Hispanic white					
Both sexes	1.0	3.3	7.5	11.3	12.6
Men	1.2	3.0	9.9	10.9	13.2
Women	0.9	3.5	5.3	11.7	12.3
Non-Hispanic black					
Both sexes	1.6	6.2	13.8	20.9	17.5
Men	1.6	5.5	13.0	16.8	14.7
Women	1.6	6.7	14.5	23.9	19.0
Mexican-American					
Both sexes	0.8	7.3	16.0	24.4	21.7
Men	0.5	5.5	15.5	18.8	17.8
Women	1.2	9.2	16.5	29.0	25.0

Four more points about patient attributes are noteworthy:

- Archimedes excludes sub-clinical phenomena that do not interest decision makers. For example, the characteristics of sarcomeres in heart muscle tissue may be interesting to heart physiologists, but not healthcare decision makers.
- Diseases are defined in terms of underlying physical attributes (signs, symptoms, and physiological attributes), as in reality. This feature allows Archimedes to reflect different definitions of disease and to incorporate new definitions.
- Diagnostic procedures, treatments, and prevention activities operate on patient attributes the same way they do in reality. For example, the drug Metformin affects the body’s production of glucose, as in reality.
- In general, the Archimedes equations that link patient attributes are not ‘laws of nature’. Rather, they are curves fitted to research data.

Percentage of the US population with diagnosed diabetes

²⁹ From the technical appendix for Eddy & Schlessinger (2003), found at www.archimedesmodel.com.

³⁰ From Harris MI et al. Prevalence of diabetes, impaired fasting glucose, and impaired glucose tolerance in US adults: the third national health and nutrition examination survey, 1988-1994. *Diabetes Care* 21:518-524, 1998. To learn more about how attributes are represented in Archimedes, see Schlessinger & Eddy (2001).

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Agent behavior rules

Patients

In Archimedes, patient behavior can include:

- seeking care for symptoms
- adhering to treatment recommendations
- taking actions that affect health-risk, such as drinking alcohol, exercising, smoking, etc.
- responding to incentives

This behavior is generally modeled with if-then behavior rules:

- if receive prescription, then fill prescription with probability p
- if fill prescription, then take drug with probability q (p and q can be functions of attributes such as age, symptom severity, co-payment levels, the number of minutes the physician spent to write the prescription, as well as random factors)
- if chest pain $>$ [pain threshold], then go to emergency room of nearest hospital (the pain threshold can be a function of age, gender, education, co-payment levels, public service announcements, and random factors)

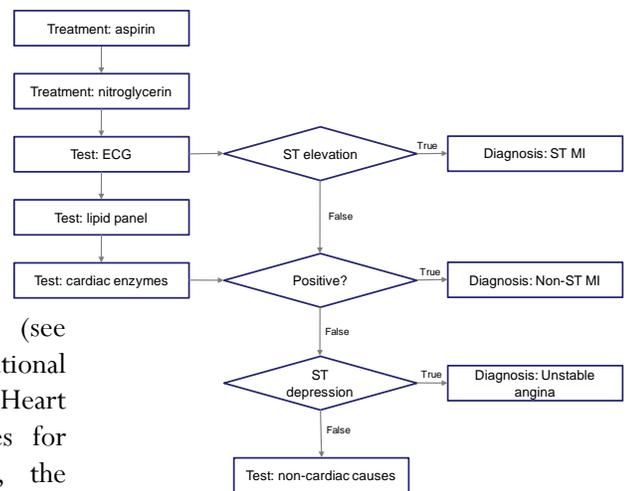
Physicians

Physician behavior can include:

- ordering diagnostic tests
- prescribing treatments
- recommending preventive procedures
- responding to incentives

Behavior is generally modeled with if-then flowcharts (see sidebar). Most of these are based on protocols from national guideline-setting organizations such as the American Heart Association, ATP III (Adult Treatment Panel guidelines for cholesterol treatment), the American Cancer Society, the American Diabetes Association, ICSI (Institute for Clinical Systems Improvement), and NGC (National Guideline Clearinghouse).

How faithfully an agent follows recommended protocols can depend on many things such as the agent’s age, medical specialty, geographic location, skill level, compensation type (salaried or fee for service), and incentives, as well as on random factors.



Simplified protocol for attending a patient with chest pain

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Agent behavior rules continued

Physicians continued

As electronic medical records are integrated with Archimedes, physician behavior can be automatically adjusted to correspond to actual current records.

Others

The behavior rules of other agents, such as hotline call operators, are generally if-then rules, based on established protocols.

Environment

The Archimedes environment can correspond to any particular geographical region. Incidence and prevalence of disease can vary by geographic region, as can the behaviors of patients and physicians.

Time

Time in Archimedes is continuous (because most biological variables vary continuously), and is organized as an ‘event queue’, a series of chronologically-ordered events. Events transpire as follows:

1. For every agent, Archimedes calculates the time of the next event, t , that will affect one of the agent’s attributes. For example, in the example given in the sidebar, t might be the time that Joe’s chest pain starts.
2. All attributes of every agent are calculated up to t .
3. Archimedes carries out the event at time t , and returns to step one.

The developers of Archimedes have worked hard to model the scheduling of patient visits and admissions as they occur in reality. For example, as in real life, Archimedes would schedule a patient’s visits for hypertension and diabetes evaluation on the same day at the same facility.

Joe’s MI

Consider an agent in an Archimedes simulation who is 40 years old and a miner. Let’s call him Joe.

On day 14,700 of his simulated life, the percent occlusion in one of Joe’s coronary arteries reaches 90 percent, and he experiences a symptom of intense chest pain. For this level of pain, his behavior rule is to drive to the nearest hospital emergency room.

Upon reaching the ER, at time 14700.60, Joe informs the staff of his chest pain. By 14700.65, the ER staff has followed its behavior rules and has given Joe aspirin, nitroglycerin, an ECG, and various blood tests.

The physician in charge, following his behavior rules, diagnoses Joe with an acute MI (myocardial infarction – a heart attack), and immediately operates. By 14700.85 Joe is resting in his hospital room with a newly-cleaned artery now occupied by a drug-eluting tube (stent) to keep the artery open.

He stays in the hospital for 4 days, and is then discharged. His behavior rules then lead him to faithfully follow the physician’s recommended follow-up treatments, giving him a significantly longer life.

Archimedes keeps a record of all of the ER staff actions, and relates them to medical codes such as Current Procedural Terminology (CPT), Relative Value Guide (RVG), and Diagnosis Related Group (DRG) codes. For each action, it also calculates the provider costs and associated health insurer reimbursements.

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

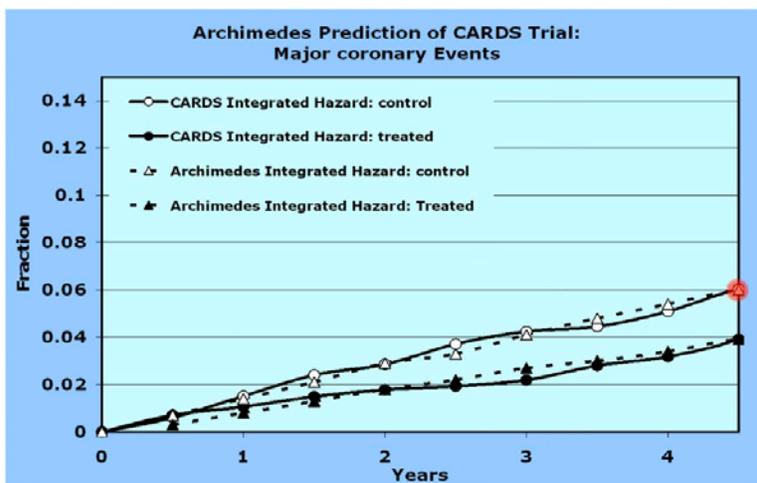
Simulation

Archimedes simulations run on a ‘farm’ of hundreds of computers, using distributed processing technology.

Validation

Archimedes has been rigorously validated against 50 clinical trials, and continues to be validated against 10-15 new trials every year.

For example, the model was prospectively validated against the CARDS (Collaborative Atorvastatin Diabetes Study), a multicenter randomized placebo-controlled study of the impact of the drug Atorvastatin on cardiovascular outcomes among people with diabetes. For the validation, Archimedes forecast results were put in a signed sealed envelope before anyone knew the study results. The following chart compares the Archimedes forecasts (dotted lines) to actual results (solid lines). Archimedes hit two of the study’s four end points on the nose.³¹



Each time the model is changed, it is automatically validated against about 20 representative studies that span the model’s capabilities. For more information about validating Archimedes, see the sidebar.

Validating Archimedes

Each equation in Archimedes is validated against independent evidence (ie, evidence not used to develop the model) when it is available, and groups of equations are validated similarly.

To find independent evidence for validation, Archimedes developers reviewed about 2000 national data sets from clinical trials and epidemiological studies (many of which are used for calibration and verification). Four main data sets for validation are:

- National health and nutrition examination survey (NHANES)
- National ambulatory medical care survey
- National hospital ambulatory care survey
- National hospital discharge survey

For validation, the following metrics are compared.

- incidence and prevalence of diagnosed conditions (for about 40 conditions)
- death rates from diagnosed conditions
- number of visits (by location and by reason)
- number of outpatient visits (by diagnosis)
- number of tests and treatments
- prevalence of prescribed interventions
- prevalence of patients taking interventions

Each metric is compared for several subpopulations:

- males, females
- age range (20-40, 40-65, 65+)
- condition (diabetes, coronary artery disease)

³¹ Trial result were reported in The Lancet 364: 685-696 (2004), and Archimedes validation results were published in an appendix to the paper Eddy AM, Schlessinger L, Kahn R (2005). *Clinical outcomes and cost-effectiveness of strategies for managing people at high risk for diabetes*. Ann Intern Med 2005; 143: 251-264. For one of the end points, stroke outcomes for people treated with Atorvastatin, Archimedes missed by a substantial margin. Atorvastatin reduced stroke incidence much more than expected from studies of other statins, and thus much more than projected by Archimedes. This result underscores what the Archimedes team itself stresses: the model can never replace controlled trials.

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Output

Common aggregate results for Archimedes include:

- disease incidence and prevalence
- symptom incidence and prevalence
- health outcomes
- utilization
- costs and reimbursements
- quality of life measures

Archimedes keeps track of every attribute that affects a patient's quality of life. For example, it tracks the time spent in every symptom and disease state, as well as the intensity of symptoms. To derive quality of life measures such as the Quality Adjusted Life Year (QALY), Archimedes applies weights from quality of life surveys.

Archimedes keeps four kinds of records:

- **true values.** It maintains the true values of all agent attributes at every time. For example, it tracks a person's true blood glucose level at every point in time (as encoded in the patient's series of blood glucose attribute functions).
- **medical record.** It maintains the results that end up in a medical record. This includes errors, because medical records are frequently inaccurate representations of truth. For example, a woman receiving a mammogram can be misdiagnosed as having breast cancer; this inaccurate diagnosis ends up in the medical record.
- **provider's knowledge.** It tracks what is in a provider agent's head. This can be different from what is in the medical record, because the provider can either misread the medical record or fail to read it (because the provider does not have access to it).
- **patient's knowledge.** It tracks what is in a patient agent's head. This is often different from both the provider's knowledge and the medical record, and can affect the patient's behavior.

It can report any value from these records, as well as any attribute value, as output.

D. PRACTICAL APPLICATIONS CONTINUED

4. Healthcare decisions continued

Application

Type 2 diabetes – one of the world’s most serious diseases – is usually non-symptomatic in its early stages. By the time people are diagnosed, problems are far advanced. Therefore, early identification is vital. In a recent research project, Archimedes was employed to discover the best screening strategy to balance cost and effectiveness.³²

Previously, no clinical trial had compared type 2 diabetes sequential screening strategies (ie, strategies that start screening at different ages and repeat at different intervals). Indeed, one result of the Archimedes project suggests that randomized clinical trials of such strategies would be unfeasible: it would require too many people and take too long to show significant difference among strategies. Archimedes compared eight screening strategies to one control (no asymptomatic screening):

Start screening	Screening interval	Stop screening
1. age 30	3 years	age 75
2. age 45	1 year	age 75
3. age 45	3 years	age 75
4. age 45	5 years	age 75
5. age 60	3 years	age 75
6. when blood pressure exceeds 140/90	1 year	age 75
7. when blood pressure exceeds 135/80	5 years	age 75
8. age 30 (maximum screening)	6 months	age 75
9. no asymptomatic screening (control)	none	never

For the simulation, Archimedes followed 325,000 patient agents from age 30 (with no type 2 diabetes) to age 80 (or earlier death).³³ When agents were diagnosed with diabetes, provider agents treated them following standard protocols. To assess the strategies, Archimedes compared their:

- incidence of type 2 diabetes
- incidence of myocardial infarction
- incidence of stroke
- incidence of microvascular complications
- quality adjusted life years (QALY)
- cost
- cost per QALY

³² R. Kahn et al. (2010)

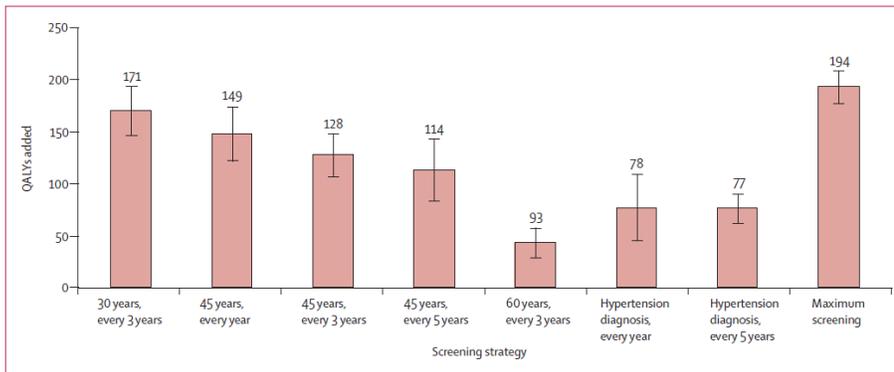
³³ Patient demographic characteristics are representative of the US population (they are based on people without diabetes in the 1999-2004 NHANES).

D. PRACTICAL APPLICATIONS CONTINUED

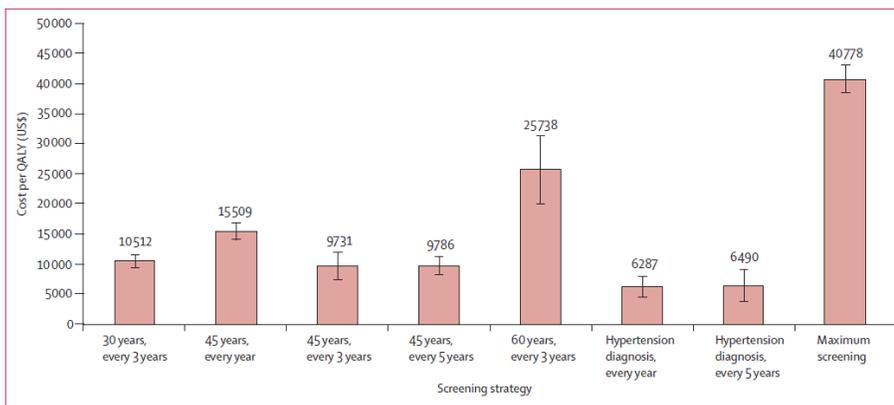
4. Healthcare decisions continued

Application continued

The charts below give two key results of the study. They show that earlier and more frequent screening increases the number of QALY, while incurring relatively modest cost per QALY. Based on these results, the project’s researchers concluded that screening for type 2 diabetes in the US population is most effective when started between the ages of 30 and 45, with screening repeated every 3-5 years.



Number of additional QALYs from screening (per 1000 people)



Cost per QALY for screening

The researchers wrote, “We believe our results are applicable to real-life settings. The Archimedes model has been constructed to be as realistic as possible and the scenarios we studied were realistic, although some simplifications were necessary.”

And the model promises to become even better (see sidebar).

ARChES

The Robert Wood Johnson Foundation (RWJF) helped solve Archimedes’s two major problems. First, because setting up and running an Archimedes simulation is extremely complicated, involving a highly-trained staff of scientists and mathematicians, it can take about six months to develop and run one analysis. The related second problem is that Archimedes is expensive: one analysis can cost hundreds of thousands of dollars.

In 2007, RWJF gave Archimedes a \$15.6 million award (the largest ever granted under its ‘Pioneer’ program) to develop a Web-based interface and delivery platform that will make the model easier and more cost-effective to use. Instead of months and hundreds of thousands of dollars, a decision maker will now need only days and tens of thousands to run an analysis. The new platform will be called ARChES.

Consistent with RWJF’s goal to transform the health and healthcare landscape, it envisions ARChES enabling organizations as varied as the Congressional Budget Office, CMS, state Medicaid offices, and the American Medical Association to apply Archimedes results to guide their decisions, thus greatly expanding the number of users and the breadth of decisions addressed.

ARChES goes live in 2012.

E. EXERCISES

1. With the Schelling segregation model, experiment with different numbers of blue and red agents, and different environment sizes. Does changing these make a difference in the model's results?
2. The Schelling segregation model in this report employs a Moore neighborhood for its behavior rule. Change the model so that it uses a von Neumann neighborhood instead (or, even better, include the neighborhood type – Moore or von Neumann – as an input parameter). Are the results different?
3. In the Schelling segregation model of this report, a dissatisfied agent moves to a random open location where it will be happy. Modify the model so that the agent moves instead to the nearest open location where it will be happy. Are the results different?
4. Modify the Schelling segregation model to include finite lifetimes for the agents. In the current model, agents live forever. As a result, once the model reaches an equilibrium state, it doesn't change. In the modified model, include two new parameters, 'minimum death age' and 'maximum death age'. Assign each agent a random death age that falls between these two parameters. When an agent is born, assign it age 0. Let it age one year for each time step ('tick'), and when it reaches its death age, allow it to die. With this change, do you see any differences in the segregation patterns?
5. Modify the Schelling segregation model to include a variety of neighbor preference percentages. Create two new parameters, a minimum preference percentage and a maximum preference percentage. Distribute random preference percentages between these two bounds to all agents. How does this change the results?
6. In chapter two of their book *Growing artificial societies*, Epstein and Axtell also explore agent migration patterns. Instead of assigning 250 agents to random locations on Sugarscape, they assign them randomly to a 20 x 20 square in the southwest corner of Sugarscape. And instead of a maximum vision of 6, they increase it to 10. Then they run the simulation to see the migration patterns that evolve as agents move around Sugarscape. Modify the Repast Symphony Sugarscape model to reproduce this simulation.

Complexity science – an introduction (and invitation) for actuaries

Chapter six: Artificial societies continued

F. TO LEARN MORE

To learn more about artificial societies, you may enjoy watching videos of Joshua Epstein describing the smallpox model and his model of how organizations adapt to market changes.³⁴ Then, you may enjoy reading one of Schelling's original books about his approach to modeling³⁵, Epstein and Axtell's book about Sugarscape³⁶, and a book about the way Complexity Science is changing the face of economics.³⁷

G. REVIEW AND A LOOK AHEAD

This chapter introduced the third of the four archetypal Complexity Science models: artificial societies. Artificial society models add an environment to the CA archetype, and so permit agents to move and interact with environmental resources.

You learned about two artificial society models, the Schelling segregation model, and Sugarscape. You also learned how such models can be constructed using Repast Symphony, and how they can be applied to simulate both an economy, an epidemic and a healthcare system.

Next, we will explore a type of Complexity Science model called 'serious games' that includes a human player. Serious games enable people to learn from simulated realities.

³⁴ J. Epstein (2008a) and J. Epstein (2008b).

³⁵ Schelling (2006)

³⁶ Epstein & Axtell (1996), one of this report's Top ten Complexity Science books.

³⁷ Beinhocker (2006), one of this report's Top ten Complexity Science books.

CHAPTER SEVEN: SERIOUS GAMES

In 1588, a thinking monk moved across the front of a room, lecturing as his students sat, watched, and listened. Occasionally he answered a question. In 1988, a thinking teacher moved across the front of a room, lecturing as his students sat, watched, and listened. Occasionally he too answered a question. In 1998, a teacher using an interactive computer game to teach students how to manage the insurance enterprise sat in the back of the room, listening and watching as her students, lost in thought, moved about the room as they interacted with their computers and with each other. Occasionally she too answered a question.

Ronald Crabb and Arnold Shapiro, 1996¹

A. INTRODUCTION

Sid Meier, a pioneer in the game industry, defined a game as ‘a series of interesting choices’. A ‘**serious game**’ is a game whose primary purpose is more than pure entertainment. Often the purpose is training, education, or discovery. Serious games are also called ‘e-learning simulations’ and ‘simulation challenges’, terms that in a corporate setting may be more readily accepted.

Serious games are a relatively new phenomenon. In 2002, the Woodrow Wilson International Center for Scholars in Washington DC started the Serious Games Initiative to encourage development of serious games addressing policy and management issues. In 2004, two additional related initiatives appeared: Games for Change, focused on social issues and social change, and Games for Health, focused on healthcare applications (see sidebar).²

This chapter introduces serious games, our fourth and final model archetype, which includes agents, agent relationships, agent behavior, an environment, and human players. You will learn about two serious game model types – the participatory model and Second life – and how serious games are applied in actuarial work.

Games for Health

Games for Health is an initiative sponsored by the Robert Wood Johnson Foundation for developing games to improve health and health care.

As part of this initiative, games have been developed to:

- Promote exercise (called ‘exergaming’)
- Change health behavior
- Provide nutrition education
- Manage disease
- Train medical personnel
- Motivate people in physical therapy
- Treat high blood pressure and depression
- Better understand the shape and function of proteins

Many companies, including CIGNA, Humana, and Kaiser, are investing in games to improve health and healthcare.

Starting in 2005, Games for Health has held an annual conference in Boston. The two-day 2010 conference included more than 400 attendees and 60 speakers.

¹ Crabb & Shapiro (1996)

² See www.seriousgames.org and ‘serious game’ at wikipedia.org.

B. BASICS

Serious games incorporate all the characteristics of the previous three model archetypes, and then add a wild card: human players.

Serious games can include from one to many players. For example, one of the model types we will explore in the next section has more than 15 million potential players, of which more than 85,000 have played simultaneously. For Complexity Science applications, the multi-player capacity of serious games is one of its most powerful features, because it allows human players to represent systems with many agents.

The most common technology for developing serious games in business environments is Adobe's Flash software. Flash technology enables the development of serious games with sound, video, animation, sophisticated interactivity, and environments in pseudo-3D.

When modern technology is combined with solid education methodologies, the result can be engaging and powerful:

- By actively participating in a realistic game, players learn by doing. They can safely try out a variety of behaviors and experience the impact not only on themselves, but also on the system as a whole.
- Multi-player serious games enable realistic social interaction, collaboration, and group problem solving. Players learn not only from the environment, but also from each other.
- Serious games can transcend barriers of language, social status, power, race, gender, and physical ability, thus drawing out the best from all players.³

Serious games appeal particularly to younger people, because they have grown up in a world where many of their social interactions are already virtual, and because they are already accustomed to using advanced technologies to play video games.

³ Purdy (2007)

C. SERIOUS GAME MODELS

There are two common serious game model types that we will explore: the participatory model and Second Life. This section describes each.

1. Participatory model

Complexity Science agent-based models need not be restricted to computers. People can also take on the agent roles and, following either simple scripted rules or their own instincts, act out the evolution of a simulation in the real world. Such a model, played without the help of computers, is called a '**participatory model**'.

Such models are relatively easy to implement, and enable participants to understand Complexity Science concepts viscerally. For example, through playing a participatory serious game, participants can learn about the often counter-intuitive emergent behaviors of complex systems.

Icosystem, a Boston company that applies Complexity Science concepts to solve business problems, often introduces people to Complexity Science using a participatory model it calls 'The Game'. To play 'The Game', ten or more participants are each asked to take on the role of an agent. Each agent is asked to randomly select two other agents from the group, agent A and agent B. Agent A is the 'protector'. Each agent is then asked to move around a room, keeping its protector between itself and agent B. The resulting movement is simply random motion without any discernable pattern.

In a second agent-based simulation, each agent becomes the protector, and is asked to move around the room, staying between agents A and B. The result is immediate, striking, emergent, and counter-intuitive: everyone clusters together in a tight knot.⁴

Another example of a real-world participatory model is the 'wave': thousands of spectators in a sports stadium following a simple rule to produce an emergent 'wave'. Another common example is real-world war games, in which soldiers fight mock battles on real terrain.

⁴ For a Java implementation of 'The Game', go to www.icosystem.com/game.htm.

Complexity science – an introduction (and invitation) for actuaries

Chapter seven: Serious games continued

C. SERIOUS GAME MODELS CONTINUED

2. Second life

Our second serious game model type is Second Life, an Internet-based pseudo-3D game where human players – represented by agents called avatars – can socialize, shop, work, learn, and explore unusual environments.⁵

The environment of Second Life is a realistic world, with geographic features similar to the earth's, including land masses, oceans, and islands. In this world agents, or avatars, can take a myriad of forms, including human, animal, plant, and robotic. Avatars can do any of the things that living objects on earth can do, such as engage in trade.

Trade in Second Life is denominated in a currency called the Linden dollar (Philip Linden is the avatar name of the game's founder, Philip Rosedale). In Second Life, Linden dollars can be used to buy, sell, or rent goods and services. They can be purchased using real international currencies, at a current exchange rate. Second Life has many banks that support Second Life economies (many of the banks recently failed because of an 'in-world' – ie, in the Second Life world – global economic meltdown).

Trade in Second Life can be profitable, in real dollars. For example, in February 2009, over 60,000 players made a profit, a few of whom grossed more than US\$ 1 million, money that can be transferred back into the real world.

In Second Life are also found cultural pursuits:

- Over 300 colleges, universities, libraries, and governments have established in-world educational institutions, teaching subjects like chemistry and foreign languages. For example, *SciLands* is devoted to science and technology education, with contributions from organizations such as NASA and NIH.
- Eight countries - Sweden, Colombia, Philippines, Albania, Estonia, Serbia, Macedonia, and The Maldives - have opened official embassies in Second Life.
- There is even a hospital and medical school in *HealthLands*.

In January 2010, Second Life had 18 million avatars, and at one time in 2009, it hosted 88,000 concurrent players.



⁵ You can explore Second Life at secondlife.com. For more information see 'Second Life' in Wikipedia.

D. PRACTICAL APPLICATIONS

This section presents four serious games relating to the work of actuaries, three about the property and casualty insurance industry, and one participatory healthcare model called CHAT.

1. Catastrophe insurance industry

In 1995, the author and mathematician John Casti became interested in the catastrophe insurance industry. He was asked to speak at a re-insurance conference exploring potential applications of science. At the conference, he was surprised to discover that most speakers focused narrowly on how science might improve the industry's ability to forecast natural catastrophes. He recalled, "I didn't believe that this was the most important problem that reinsurance would be facing or that science could shed some light on. Rather, I felt that a much more interesting and important question was: 'How do you understand your place as a firm within the overall industry?'"

From that experience and his work with the Santa Fe Institute, Casti conceived the idea to develop an agent-based model of the catastrophe insurance industry. To fund the model's development and to assemble subject-matter experts, he organized a consortium of fifteen companies that included Ernst & Young, Swiss Re, ItalRe, Winterthur, Marsh & McLennan, Los Alamos National Laboratory, and the Santa Fe Institute. Work on the model, named 'Insurance World', began in 1997 and was largely completed in 2002.⁶

Although its developers did not advertise it as such, the model is in the form of a serious game. It enables its players (in the role of insurance and re-insurance company executives) to devise business strategies to react to varying environmental and competitive conditions, and then witness the systemic results of such strategies. The environmental conditions include random natural catastrophes and terrorism, changing economic conditions, and different consumer markets. As the game progresses, players can monitor the impact of their strategic choices through detailed quarterly financial statements and market share results, and can change their strategies. By the end of the game some companies have thrived, and some have failed.

⁶ For information about Insurance World and its origins, see Segre-Tossani (2003) and Gionta (2000).

D. PRACTICAL APPLICATIONS CONTINUED

1. Catastrophe insurance industry continued

Insurance World consists of ten agents, a three-tiered environment, and players. Following is a description of these components. Because Insurance World was a commercial product, details about its operation were not published. Consequently, the following description is only an overview of the model.

Agents

The agents are five insurance companies and five re-insurance companies, the attributes of which include:

- strategic goals, such as:
 - growth rate
 - cost of capital
 - net combined ratio (ratio of annual retained losses – ie, losses not re-insured – plus costs, divided by retained premiums)
 - ratio of premiums to total assets
 - ratio of premium reserves to the sum of annual retained losses plus costs
 - efficiency of capital use (subscribed capital/total assets)
 - portfolio diversification
 - level of investment in catastrophe bonds
- market share
- balance sheet items such as:
 - current assets
 - fixed assets
 - subscribed capital
 - premium reserves
 - outstanding loss reserves
 - debt
- earnings statement items such as:
 - retained premiums
 - retained losses
 - costs
 - interest income

The agents are related through their business interactions. For example, each of the five insurance companies might re-insure its risks with two of the five re-insurance companies.

Complexity science – an introduction (and invitation) for actuaries

Chapter seven: Serious games continued

D. PRACTICAL APPLICATIONS CONTINUED

1. Catastrophe insurance industry continued

Environment

In Insurance World, the environment consists of three sub-environments with which the agents interact:

- **consumer market regions:** The international geographic regions in which the agents operate. Up to 10 market regions can be defined, including regions of the US, Europe, and Japan.
- **economy:** The economy in which the agents do business, with attributes including inflation rates, as well as quarterly values of short term bonds, long term bonds, three stock markets, real estate, and catastrophe bonds.
- **nature:** The source of natural catastrophes, including earthquakes, windstorms (such as hurricanes), and floods. Terrorism is also included in this sub-environment.

Agent behavior

Agents act in order to attain their strategic goals within the constraints of competition and solvency requirements. Their behavior includes receiving premiums, paying out losses, investing in the economy's financial markets, and re-insuring risks, all in accord with their strategic goals. To re-insure risks, the insurance companies negotiate with re-insurance companies to determine the amount of risk that is ceded and its cost. Insurance companies with larger market shares wield greater clout in these negotiations.

D. PRACTICAL APPLICATIONS CONTINUED

1. Catastrophe insurance industry continued

Environment behavior

In the model, each consumer market interacts with the agents through its solvency requirements, and its preferences for different types of insurance companies (some markets choose insurance companies based mainly on brand recognition, and some choose on the basis of price).

The economy produces random fluctuations in financial market values and inflation rates.

Nature generates random catastrophic events – such as floods, earthquakes, windstorms (hurricanes), and terrorist attacks – in each of the model’s markets.

Time

Insurance World simulations run for ten years, in quarterly time increments.

Players

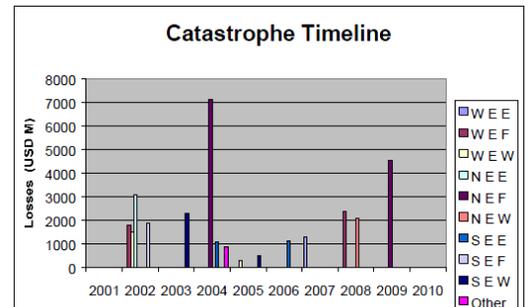
The players define each agent’s strategy. As conditions change during a simulation, the players can revise their strategies.

Output

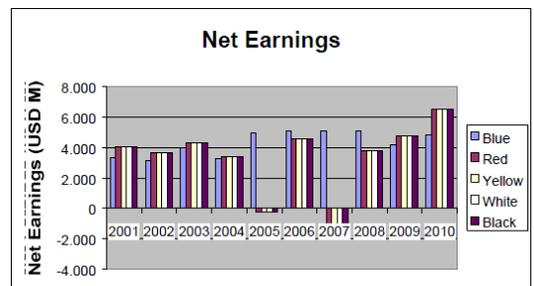
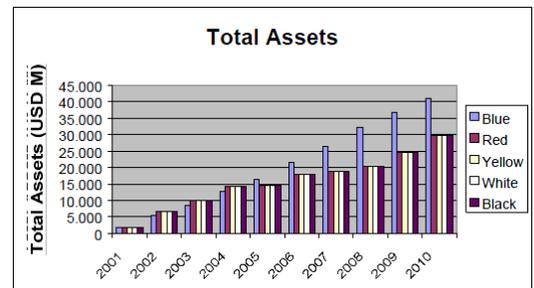
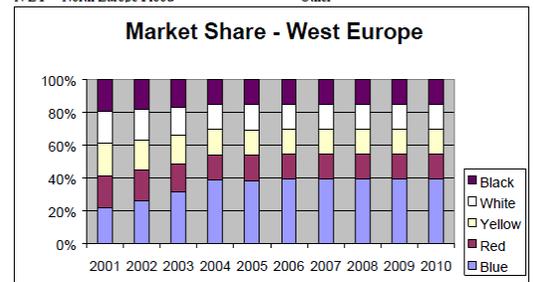
The model produces the following output (see charts at right)⁷:

- The incidence and magnitude of catastrophic events
- Each agent’s balance sheet, earnings statement, and financial ratios
- Each agent’s market share in each consumer market
- The values of each financial instrument (bonds, stock markets, etc.)

In 2003, because of an argument about ownership rights, the model was withdrawn from the market. The only company now using it appears to be Swiss Re, in its risk management group.



WEE = West Europe Earthquake
 WEF = West Europe Flood
 WEW = West Europe Windstorm
 NEE = North Europe Earthquake
 NEF = North Europe Flood
 NEW = North Europe Windstorm
 SEE = South Europe Earthquake
 SEF = South Europe Flood
 SEW = South Europe Windstorm
 Other



⁷ The charts are from Gionta (2000).

D. PRACTICAL APPLICATIONS CONTINUED

2. P&C re-insurance market price dynamics

The US property and casualty (P&C) re-insurance market is unusual: It has few major players (less than ten), almost no product differentiation, and prices that are highly cyclical (with a cycle time of about eight years). In 2005, Jens Alkemper and Don Mango (an actuary) developed an agent-based model to explore why its prices are cyclical.⁸

Following is a description of the model's characteristics. However, because the model is not publicly available and its published description is not complete, the following is only an overview:

Agents

The model includes three agents. Each is a P&C re-insurance company, with the following attributes:

- a portfolio, consisting of multiple books of business, each with a number of exposure units, a price per exposure unit, and expected claims per exposure unit. It is assumed that claims are paid out over four years.
- premium revenue (equal to exposure units * price per unit)
- ultimate claim payments (equal to exposure units * expected claims per exposure unit)
- reserve liability (equal to the total ultimate claim payments minus the cumulative amount of claim payments to date)
- expenses
- liability (equal to the sum of reserve liabilities for all books of business)
- assets (equal to accumulated premiums for all books of business less expenses and claim payments)
- reinsurance capacity available
- capital (equal to assets minus liabilities)
- required capital (implemented as factors multiplied by exposure units)
- constraint ratio (a percentage that constrains the agent's ratio of capital to required capital)

At the start of the simulation, each agent is given an initial book of business and an amount of assets. Each agent is related to the others, because they all participate in one competitive market.

⁸ Alkemper & Mango (2005)

D. PRACTICAL APPLICATIONS CONTINUED

2. P&C re-insurance market dynamics continued

Agent behavior

At each time step, based on its capital, required capital, and constraint ratio, an agent decides the additional number of exposure units it will underwrite, and bids that number.

Environment

The environment is the market in which the agents operate, and a source of catastrophes.

Environment behavior

Based on a simple demand curve (see the chart at right), at each time step the environment computes the market price as a function of the agents' total bids.

The environment also initiates a catastrophe at time step 60. The catastrophe absorbs approximately 20 percent of each agent's capital.

Time steps

Each time step is a year, and the model is allowed to run for 60 time steps for the system to reach equilibrium. The actual simulation is then from time step 60 to time step 80.

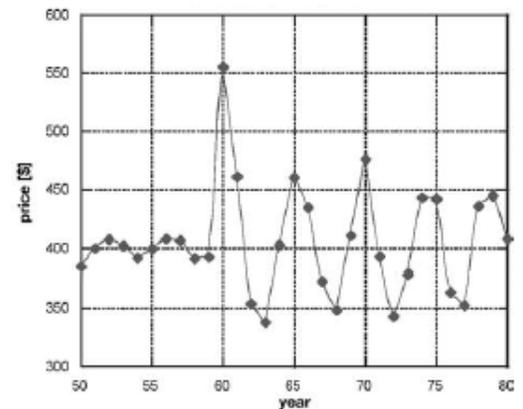
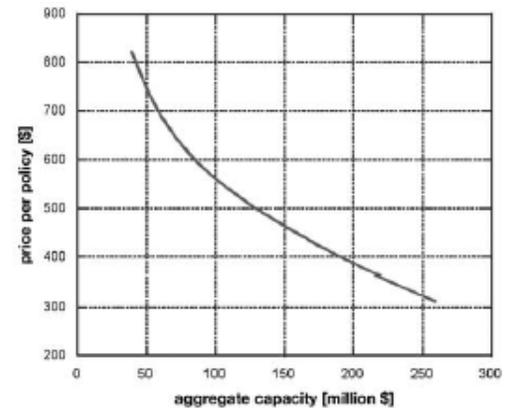
Output

The primary output is the market price.

Players

Each player assumes the role of an agent. Based on historical prices, and an agent's historical financial results, at each time step a player submits a blinded bid for exposure units.

As you can see from the chart at right, even with the model's simple behavior rules, it produces dramatic price cycles with a cycle time of about five years. This result is robust: The cycle time does not depend on the time period over which claims are paid for a given book of business, nor on the shape of the demand curve. The critical and counter-intuitive insight from the model is that the price cycles are an emergent property of the re-insurance system itself, apparently arising from the bidding strategies.⁹



⁹ This conclusion and the charts are from Alkemper & Mango (2005).

D. PRACTICAL APPLICATIONS CONTINUED

3. Managing an insurance company

In the early 1990s, Ronald Crabb and Arnold Shapiro (an actuary) developed a serious game to teach people how to manage a company that insures automobiles. Following is a description of its characteristics. However, because the model is not publicly available and its published description is not complete, the following is only an overview:¹⁰

Agents

There are four agents, each of which is an automobile insurer. Its attributes include:

- strategic variables:
 - advertising budget
 - premium price for each consumer risk level
 - commission rate
 - education and training budget
 - claims policy
 - underwriting policy
 - percentage allocation of assets to cash, short-term investments, high-quality bonds, low-quality bonds, and common stock
 - level of desired risk for investing in common stock
- assets
- liabilities
- expenses (the underwriting expense ratio is assumed to be 29 percent, the combined ratio is assumed to be 111 percent, and the loss and loss adjustment expense ratio is assumed to be 82 percent)
- surplus
- unearned premium reserve account
- loss reserves account
- adjusted surplus (statutory surplus adjusted to reflect the equity in the unearned premium reserve account and in the loss reserves account)

At the start of each simulation, each agent has the same allocation of policies, assets, liabilities, and surplus.

Because they are competitors in one market, the agents are all related to one another.

¹⁰ Crabb & Shapiro (1996)

D. PRACTICAL APPLICATIONS CONTINUED

3. Managing an insurance company continued

Environment

The environment consists of a consumer marketplace, and a financial marketplace. The consumer marketplace has a constant number of consumers with three levels of risk: standard, substandard, and preferred. The financial marketplace consists of short-term investments, high-quality bonds, low-quality bonds, and common stock.

Agent behavior

At the start of the simulation, each agent establishes its company strategic policy by assigning a value to each strategic variable. It then writes policies, collects premiums, pays claims, and invests surplus. Policies are assumed to be written uniformly throughout each time step, and premiums paid semiannually.

Environment behavior

The environment determines the rates of return on financial instruments, each agent's market share, and the incidence of claims. Less risky investments have lower rates of return and lower volatility. And the claims tail is limited to three years.

The market share of the i -th insurer in the j -th consumer market risk level (standard, sub-standard, and premium) at time step t is:

$$MS_{i,j,t} \approx [AW_{i,t}][PW_{i,j,t}][CW_{i,t}][EW_{i,t}][CPW_{i,t}][UPW_{i,j,t}]$$

where:

$[AW_{i,t}]$ is the advertising weight

$[PW_{i,j,t}]$ is the price weight

$[CW_{i,t}]$ is the commission weight

$[EW_{i,t}]$ is the education and training weight

$[CPW_{i,t}]$ is the claims policy weight

$[UPW_{i,j,t}]$ is the underwriting policy weight

The weights are functions of the agent strategic variables. For example, the price weight is larger for an agent that sets a lower price relative to competitors, and thus increases that agent's market share.

D. PRACTICAL APPLICATIONS CONTINUED

3. Managing an insurance company continued

Environment behavior continued

Similarly, relatively larger commissions increase market share (but increase the underwriting expense ratio), a looser underwriting policy increases market share (but increases the loss and loss adjustment expense ratio), a looser claim paying policy increases market share (but increases the loss and loss adjustment expense ratio), a larger advertising budget increases market share (but increases the underwriting expense ratio), a larger education and training budget increases market share (but increases the underwriting expense ratio and decreases the loss and loss adjustment ratio)

As in the real world, the impact of variables on market share is lagged, in order to avoid a massive shift in market share in one time step.

Time steps

Each time step is one year. The number of time steps included in a simulation is not reported.

Output

At each time step, simulation output includes the following for each company:

- adjusted surplus
- loss ratio
- expense ratio
- investment earnings

This output is presented graphically.

Players

Four human players assume the roles of the four agents. Each player establishes an agent's initial strategy by assigning a value to each strategic variable.

The player whose agent has the largest increase in its adjusted surplus is the winner.

D. PRACTICAL APPLICATIONS CONTINUED

4. CHAT

Choosing health plans all together (CHAT) is a participatory agent-based model in which a group of agents designs its own health insurance plan. It was developed in 1995 by physicians at the University of Michigan Medical School and the National Institutes of Health. The purpose of the model is to discover the health insurance plan that emerges from the bottom up, when each member of a community is involved in its design.¹¹ The CHAT model was demonstrated for actuaries at the Health Spring Meetings in 2008 and 2009. Following are its characteristics:¹²

Agents

An agent is a person in a group that is deciding what healthcare benefit types (eg, hospital services, physician services, tests, drugs, home health care, dental care, etc.) and levels of coverage (usually three) to include in its health insurance plan. To purchase their plan, the group has a limited budget.

Environment

The environment includes the healthcare system of the community in which the agents live, health events that befall the agents, and actuaries to calculate the cost of benefits that the group chooses.

Environment behavior

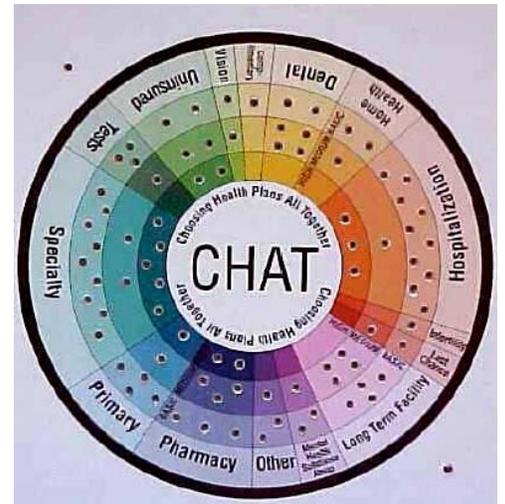
At each time step, the environment provides a random health event for each agent, to enable the agent to test the reasonability of its budget allocation. Also, actuaries in the environment calculate the cost of each healthcare benefit, and the cost of each total health insurance plan that agents choose.

Time steps

Generally, there are four time steps.

Agent behavior

At each time step, agents allocate the group's budget among the healthcare benefit choices. Generally, in the first time step each agent individually decides the allocation; in the second step, agents in groups of three decide the allocation; in the third, the whole group decides; and in the fourth, each agent decides again individually.



¹¹ For an interesting video about the application of CHAT in rural India, see Microinsurance (2008).

¹² Danis (2003)

Complexity science – an introduction (and invitation) for actuaries

Chapter seven: Serious games continued

E. EXERCISES

1. How would you recreate the Schelling segregation model using a participatory model type?
2. Create an avatar and pay a visit to SciLands and HealthLands in Second Life. What facilities might you establish in Second Life to appeal to actuaries? How would you implement and monitor a complex adaptive system in Second Life?
3. Identify an aspect of your work that you would like to explore with a serious game, and then design such a game. Describe the agents, agent relationships, agent behavior, environment, environment behavior, and players.

F. TO LEARN MORE

To learn more about serious games, you may enjoy watching a video about the Serious Games Institute,¹³ or reading the book titled, *The complete guide to simulations & serious games*.¹⁴

G. REVIEW AND A LOOK AHEAD

This chapter introduced the fourth of the four archetypal Complexity Science models: serious games. Serious game models add human players to the artificial society archetype, and so enable people to interact with virtual environments and each other to explore the workings of complex systems.

You learned about two serious game models, the participatory model and Second Life. You also learned how such models can be applied to simulate aspects of the property and casualty insurance industry, as well as enable a community to design an appropriate health insurance plan.

This chapter completes Part II of this report. In Part III, we will explore how actuaries can use Complexity Science to improve their work.

¹³ Serious Games Institute (2007)

¹⁴ Aldrich (2009)

PART III: AN INVITATION

Like most people, actuaries prefer staying in their ‘comfort zones,’ or areas of familiarity, particularly when it comes to research, where most of our efforts are spent digging further into known models and topics. This is important, as a science needs to refine and extend its core knowledge base. But the tendency to stay close to home is fine for the health of a science only as long as the underlying environment itself is relatively stable.

Unfortunately, we are in an unstable, complex, evolving environment. Our employers face serious problems that span insurance, finance, the capital markets, and the economy as a whole. These problems reach across multiple professions—many comfort zones—and thus have no ‘owner profession.’ Being the professionals who, by our own proclamation, ‘make financial sense of the future,’ it is incumbent upon us to step up and play a leadership role in formulating research solutions to these problems.

Shaun Wang and Donald Mango (actuaries)¹

¹ Wang & Mango (2003)

CHAPTER EIGHT: THE COMPLEX SYSTEMS ACTUARY

*While Complexity Economics strips away our illusions of control over our economic fate, it also hands us a lever – a lever that we have always possessed but never fully appreciated. **We may not be able to predict or direct economic evolution, but we can design our institutions and societies to be better or worse evolvers.***

Eric Beinhocker¹

A. INTRODUCTION

In this chapter, I invite you to change the way you work.

Actuaries tend to work in special niches. You are a pension actuary, a reinsurance actuary, a property-casualty actuary, a life insurance actuary, or a health actuary. And there are sub-specialties: you are a reserving health actuary or a trend health actuary or a pricing health actuary. Each niche has its special actuarial methods, handed down from actuary to actuary, guild-like, over decades, with actuaries from one guild often understanding little of the others.

Yet, today's great problems do not lie in niches. Today's challenges are systemic: How do we ensure the viability of pension systems? How do we design an affordable health system? How do we prevent the collapse of financial systems?

When niche-ensconced actuaries cannot address such systemic societal problems, society finds solutions on its own. The result? The solutions are often suboptimal, and actuaries disappear. This happened in the US in the 1970s when actuaries did not address systemic pension problems. Society found a solution – ERISA – and pension actuaries disappeared or were marginalized to implement Byzantine regulations. It happened again in the US in 2010: for decades health actuaries fussed with their reserves and trends and pricing, unable to address systemic healthcare problems. Society found its own solution – healthcare reform – and soon, just as in the UK and Canada, US health actuaries may disappear or be marginalized.

As you have learned, all actuaries work in complex systems, and all complex systems have common characteristics. This chapter invites you to step out of your niche, to become what I call a **'complex systems actuary'**.

¹ Beinhocker (2006), page 324.

Complexity science – an introduction (and invitation) for actuaries

Chapter eight: The complex systems actuary continued

A. INTRODUCTION CONTINUED

A complex systems actuary is a professional who addresses problems in complex systems of all types, not just pension, insurance, and healthcare systems, but also financial systems, city and state systems, and corporate systems – any complex system where people, money, and contingency intersect. A complex systems actuary is a professional who finds potential solutions to systemic problems ranging from efficiency and allocation issues, to issues involving strategic risk management, a professional who effectively communicates the potential solutions and helps to change social policy.

This chapter describes complex systems actuaries: their work, clients, employers, knowledge and skills, competitors, and social impact. At the chapter's end are exercises to sharpen your understanding.

B. WORK

Complex systems actuaries improve how complex social systems work. Such systems range from small businesses to multi-national conglomerates, from local investment clubs to international banking systems, from small cities to multi-state regions.

To do their work, complex systems actuaries:

- **Define the system.** Using Complexity Science tools, the actuary determines the system's agents, relationships, behavior rules, and environment.
- **Construct an agent-based model.** The actuary builds an agent-based model to simulate the system and its problem.
- **Generate potential solutions.** By varying system components (agents, relationships, behavior rules) the actuary finds combinations that promise to solve the problem. The actuary may incorporate learning algorithms in the model to efficiently sort through the design space and improve the system's location on its fitness landscape.²
- **Communicate the results.** Using the agent-based model, the actuary communicates potential solutions to stakeholders. To facilitate communication, the actuary may construct a serious game for the stakeholders to play. A key advantage of agent-based models is that they facilitate communication.
- **Implement a solution.** The actuary helps to implement the solution stakeholders choose, and then monitors the results.

² To review the concepts of 'design space' and 'fitness landscape', see Chapter one.

Complexity science – an introduction (and invitation) for actuaries

Chapter eight: The complex systems actuary continued

B. WORK CONTINUED

For example, suppose a state approaches you, a complex systems actuary, to solve its problem of unsustainable healthcare expenditure increases. What do you do?

You first define the system's agents (its physicians, patients, hospitals, medical suppliers, insurance companies, etc.), and their relationships, behavior rules, and environment. To define the behavior rules, you search the literature to find what is known, and then you may perform interviews or even controlled behavioral experiments to get the rest (see Chapter two).

Perhaps using Repast, you construct an agent-based model to simulate the system and its problem. Perhaps you even construct your model as a serious game, allowing humans to play agent roles. Next, you and others play with the model, to find new behaviors and relationships that solve the problem. You present potential solutions to state stakeholders, and encourage them to also play with the model, see solutions for themselves, and choose a preferred solution. To help stakeholders sell their chosen solution, you may even put the model on a public website and encourage the state's inhabitants to play with it. Finally, you help implement the chosen solution, and monitor its effects.

Other potential clients, and problems they might bring you, are:

Business firms (including insurance organizations)

- Find the causes of business cycles, and ways to dampen them (see sidebar)
- Find business and investment strategies that are robust
- Determine organizational structures that lead to long-term fitness
- Determine the potential impact of new products or policies
- Find sources of supply and demand fluctuation and ways to manage them
- Test validity of business assumptions about the environment
- Identify synergistic opportunities among stakeholders
- Determine human resources required to address market needs
- Design strategies to manage enterprise risk
- Determine the causes of certain fat-tailed risks, and ways to manage them
- Identify potential loss from specified emerging risks

Great opportunity for collaboration

“Scientists are realizing that collective systems exhibit interaction effects that cannot be predicted from the behavior of the individual elements. J. Doyne Farmer of Santa Fe Institute has published several papers showing that capital market price volatility can in part be explained by the interaction of competing trading strategies among different types of traders. ... Avinash Persaud of Gresham College has written repeatedly on the paradoxical impact of market-sensitive risk management policies on banking system stability. His premise, ‘Introduction of market-based risk based capital requirements leads to uniformity of risk appetite among participants, and therefore uniformity of response to market volatility.’ The uniformity compounds and reinforces itself as participants react to each others' reactions, leading to market crises. Applying similar logic to the insurance market, it is likely that strategic interaction plays a material role in insurance pricing cycles.

When studying such systems, researchers must be wary not to extrapolate incorrectly from local conditions. These systems are non-linear, and the strategy assessment is multilateral not unilateral. Theories and policy recommendations may only be testable using nontraditional scientific approaches and media. Examples include simulated economies and agent-based models. Possible worthwhile research goals include strategy robustness testing - which plan works best, factoring in all the possible things others could do? - and policy recommendations - how should regulators monitor and control the system to maximize stability?. This is a great opportunity for collaboration outside insurance and retirement systems.”

Don Mango (an actuary)³

³ Mango (2005)

Complexity science – an introduction (and invitation) for actuaries

Chapter eight: The complex systems actuary continued

B. WORK CONTINUED

Regulatory bodies

- Determine the potential impact of proposed regulations and policy changes

Health insurers

- Find the drivers of healthcare expenditure trends and a way to manage them
- Determine robust strategies for provider negotiations
- Determine the potential impact of pricing, reimbursement, and policy changes

States and countries

- Determine the potential impact of policy changes
- Design a social security policy that works
- Design a healthcare system that works

C. CLIENTS

Potential clients for complex systems actuaries include:

- business firms, to help with any of their systemic problems involving people, money, and contingency; not just traditional actuarial problems.
- cities, states, countries, and international regions
- regulatory bodies, such as the NAIC
- insurance and reinsurance sectors
- financial sectors

This report has described many examples where Complexity Science has been employed to solve problems for all these entities.

D. EMPLOYERS

The complex systems actuary might work for:

- a think tank like the Brookings Institution
- a research institute like the Santa Fe Institute
- a consulting firm
- a large business enterprise
- an insurance or reinsurance company
- a government entity

E. KNOWLEDGE AND SKILLS

Complex systems actuaries apply the following knowledge and skills:

- Complexity Science concepts and tools
- agent-based modeling
- actuarial concepts, methods, and tools
- risk management concepts and tools
- general business knowledge
- communication skills

Although many actuarial concepts and tools (such as time value of money, contingencies, and financial concepts) are useful to complex systems actuaries, some get in the way: Most actuarial models are based on projection of aggregate historical data patterns, and provide no insight into behavior rules or agent relationships. To simulate complex systems, they are not useful (see sidebar). Complex systems actuaries will not need most traditional actuarial methods; instead they will use bottom-up agent-based methods.

But complex systems actuaries will need excellent communication skills, because changing social policy requires effective communication with a variety of stakeholders. Traditional actuaries have found communication to be difficult, partly because actuarial concepts are often abstruse, rigid, and remote from business reality, but also because actuaries do not receive adequate training in communication arts. Complexity Science makes communication easier: its results are more visual and intuitive, and people easily relate to agents, relationships, and behavior rules.

F. COMPETITORS

Competitors are few, because few have the combination of knowledge and skills necessary to solve complex social system problems. For example:

- Economists and management consultants generally lack the knowledge of Complexity Science and the ability to construct agent-based models.
- Complexity scientists generally lack the business knowledge of actuaries.

Bees on average

In their book *Complex adaptive systems*, Miller and Page describe how bees keep their hives at constant temperature:

“For honey bees to reproduce and grow, they must maintain the temperature of their hive in a fairly narrow range ... When the hive gets too cold, bees huddle together, buzz their wings, and heat it up. When the hive gets too hot, bees spread out, fan their wings, and cool things down.

Each individual bee’s temperature thresholds for huddling and fanning are tied to a genetically linked trait. Thus, genetically similar bees all feel a chill at the same temperature and begin to huddle; similarly, they also overheat at the same temperature and spread out and fan in response.

...

Hives with genetic diversity produce much more stable internal temperatures. As the temperature drops, only a few bees react and huddle together, slowly bringing up the temperature. If the temperature continues to fall, a few more bees join into the mass to help out. A similar effect happens when the hive begins to overheat. ...

In this example, considering the average behavior of the bees is very misleading. The hive that lacked genetic diversity – essentially a hive of averages – behaves in a very different way than the diverse hive. Here, average behavior leads to wide temperature fluctuations whereas heterogeneous behavior leads to stability. To understand this phenomenon, we need to view the hive as a complex adaptive system and not as a collection of individual bees whose differences cancel out one another.”⁴

⁴ Miller & Page (2007), page 15.

Complexity science – an introduction (and invitation) for actuaries

Chapter eight: The complex systems actuary continued

G. SOCIAL IMPACT

The potential social impact of complex systems actuaries is great. For example, had complex systems actuaries constructed models of the US healthcare system, and identified solutions to its expenditure and coverage problems, the healthcare reform debate could have been more rigorous and fact-based, less emotional and political.

Some hold that a model of the extremely complex US healthcare system is impossible to construct and validate. I, and others, disagree.⁵

H. EXERCISES

1. Consider a real-world complex system problem, and describe how you, as a complex systems actuary, would solve it and help to change social policy.
2. Think of an important complex system problem of your current employer. Prepare a proposal to solve the problem, using the knowledge and skills of a complex systems actuary. Present the proposal to your manager and obtain feedback. (And please let me know if you carry out the project.)

I. TO LEARN MORE

To learn more about communicating the results of complex system simulations, see works by Edward Tufte, such as *The visual display of quantitative information*, *Beautiful evidence*, and *The cognitive style of Power Point: pitching out corrupts within*.⁶ You may also enjoy the paper titled *Design guidelines for agent-based model visualization*.⁷

J. REVIEW AND A LOOK AHEAD

This chapter presented a vision of the ‘complex systems actuary’, including this new professional’s work, clients, employers, knowledge and skills, competitors, and potential social impact. The next chapter presents a plan to develop such actuaries.

⁵ For example, see Strip, Backus, Strickland, & Schoenwald (2005).

⁶ Tufte (2001), Tufte (2006a), and Tufte (2006b).

⁷ Kornhauser, Wilensky, & Rand (2009)

CHAPTER NINE: NEXT STEPS

Progress will come not by refining existing models, but by breaking barriers and incorporating new classes of models and their associated estimation and checking methods.

James Hickman (actuary)²

This short chapter suggests several steps to nurture the development of complex systems actuaries:

1. Form a development oversight team

Assemble a group of interested actuaries to flesh out these steps, and oversee their implementation.

2. Implement a discussion forum

Implement a wiki, LinkedIn group, email discussion list, or other mechanism for actuaries interested in Complexity Science to exchange ideas.

3. Hold Complexity Science workshops

Hold a series of workshops at Society of Actuary (SOA) and Casualty Actuary Society (CAS) meetings about applications of Complexity Science in actuarial work.

4. Promote Complexity Science research

Encourage SOA and CAS special interest sections to fund basic Complexity Science research, promote agent-based models as a research product (see sidebar), and offer prizes for research that demonstrates how Complexity Science more effectively addresses the problems actuaries face today, and expands the scope of problems they can face tomorrow.

5. Establish outside relationships

Establish relationships with institutes that perform Complexity Science research, such as

- Santa Fe Institute
- Brookings Institution
- New England Complex Systems institute
- Center for the Study of Complex Systems (University of Michigan)
- Center for Social Complexity (George Mason University)
- Institute on Complex Systems (Northwestern University)

For example, actuaries can become visiting scholars at these institutions, or co-author articles with complexity scientists.

The agents are coming

“If you have seen the enormously popular software *The Sims*, you were probably intrigued, but you may not have realized that you were looking at part of a scientific revolution involving adaptive agents. Adaptive agents are software entities that, when placed in a computer environment, monitor the state of that environment, and, armed with rules of behavior, interact with it. In the case of *The Sims*, the agents are people of various types, placed in a house or hotel or city, interacting in ways both mundane and hilarious. One of the key elements to the popularity of *The Sims* is its complete unpredictability. There is literally no way to predetermine the aggregate result of placing a certain set of Sims in a certain environment with certain initial conditions. The only way to find out what will happen is by *watching the situation unfold*.

The adaptive agent paradigm (AAP) is gaining momentum in many areas of science: financial markets, water policy, organizational and network theory, and the military. ...

What might this mean for our science? Here are some speculations:

1. Forced group interaction (similar to multiplayer gaming): the days of the isolated practitioner are limited.
2. **Software as research product: software will have formal recognition as a communication medium. ...**
3. Policy analysis: regulators could use this as a means of testing the impact of policy changes (eg, fair value accounting).
4. Test impact of changes: allow testing of the aggregate effects of changes in rates, class plans, laws, or agent compensation in order to determine beforehand the likely impact of changes.”¹

Donald Mango (actuary)

¹ Mango (2004)

² Hickman (1997)

Complexity science – an introduction for actuaries

Chapter nine: Next steps continued

6. *Include Complexity Science in actuarial education*

Include the concepts and tools of Complexity Science on the actuarial exams, perhaps as part of an ASA module.

7. *Establish a Complexity Science special interest section*

Because the work of complex systems actuaries cuts across all traditional actuarial areas, establish a new SOA special interest section devoted to Complexity Science.

8. *Establish a complex systems actuarial research institute*

Establish a complex systems actuarial research institute, similar to the Santa Fe Institute. The institute should:

- Maintain an agent-based model archive, where actuaries can view and share models, documentation, and supplemental files. As you have seen, actuaries have made sporadic efforts to use agent-based models to solve complex system problems. But, there has been no cumulative progress. By maintaining a model archive, the Institute will foster cumulative progress (see sidebar).
- Develop and maintain tools and components to help actuaries construct agent-based models.
- Maintain large databases relating to complex systems of interest to actuaries, such as claims data of health insurers.
- Perform controlled experiments to better understand the behavioral rules of agents in systems of interest to actuaries.
- Recommend agent-based modeling standards, and standard evaluation criteria for agent-based models. (see sidebar)
- Perform research to deepen our understanding of complex systems of interest to actuaries, as well as advance our methods for simulating such systems.
- Solve complex system problems for specific organizations, and thereby help to develop social policy.

The institute, like the Santa Fe Institute, will obtain seed money from organizations that stand to benefit from its work. Then its funding will come from grants, foundations, and project fees.

*To take these steps,
I request the pleasure of your company.*

RSVP

The tortoise and hare

Takadama and Shimohara recommend the following to foster cumulative progress in agent-based modeling:

“(a) **Common test-beds:** Sharing common test-beds is a promising approach for cumulative progress. The reasons are summarized as follows: (1) common test-beds enable researchers to narrow an argument down to concrete and detailed issues, which help to provide a fruitful and productive discussion; and (2) common test-beds encourage researchers to share results, which can lead to progress in the field by comparing results or competing with other researchers.

(b) **Standard computational models.** Standard computational models are necessary for cumulative progress. With them (1) researchers do not need to design computational models, and this can contribute to bringing several researchers together and led to progress in the field; and (2) common parts of various research efforts can become clear through the development of libraries of computational models, providing the essential parts of agent-based simulations.

(c) **Validation and advance of older work.** It is important to validate older results and advance older work for cumulative progress. In this case, the replication of older models is essential to validate and advance older work. To promote this, researchers should share and understand what things have already been done and what things have not in an agent-based approach.

(d) **Standard evaluation criteria:** Standard evaluation criteria for results ... are indispensable for cumulative progress. Although it is difficult to evaluate results appropriately, it is important to apply the same evaluation criteria.”

Like the tortoise and hare, “continuous small progress is more important than intermittent rapid progress.”³

³ Takadama & Shimohara (2002)

TOP TEN COMPLEXITY SCIENCE BOOKS

A. INTRODUCTION

This chapter describes what are, in my opinion, the top ten books to help you learn more about Complexity Science. I selected these books based on the following criteria:

- Together they should cover four key areas of Complexity Science – complex adaptive systems, networks, behavioral economics, and agent-based modeling – as well as the four archetypal models of Complexity Science.
- Each book should be introductory in nature, yet contain sufficient new material to help actuaries apply the insights and tools of Complexity Science.
- Each book should be more useful for actuaries to read than other similar books.
- The books should be easily accessible, either through www.Amazon.com or a library.

In the order in which I suggest that you read them, they are:

1. *Complexity: the emerging science at the edge of order and chaos*
by M. Mitchell Waldrop (1992)
2. *Managing business complexity: discovering strategic solutions with agent-based modeling and simulation*
by Michael J. North and Charles M. Macal (2007)
3. *Linked: how everything is connected to everything else and what it means for business, science, and everyday life*
by Albert-László Barabási (2003)
4. *The structure and dynamics of networks*
by Mark Newman, Albert-László Barabási, and Duncan Watts (2006)
5. *A new kind of science*
by Stephen Wolfram (2002)
6. *Predictably irrational: the hidden forces that shape our decisions*
by Dan Ariely (2008)
7. *Complex adaptive systems: an introduction to computational models of social life*
by John M. Miller and Scott E. Page (2007)
8. *Growing artificial societies: social science from the bottom up*
by Joshua M. Epstein and Robert Axtell (1996)
9. *Generative social science: studies in agent-based computational modeling*
by Joshua M. Epstein (2006)
10. *The origin of wealth: evolution, complexity, and the radical remaking of economics*
by Eric D. Beinhocker (2006)

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

1. COMPLEXITY

THE EMERGING SCIENCE AT THE EDGE OF ORDER AND CHAOS

BY: M. MITCHELL WALDROP (1992)

Structure: 9 chapters, 380 pages, with index and bibliography

Recommended reading: all chapters

Themes:

complex adaptive systems..	<input checked="" type="checkbox"/>	behavioral economics...	<input checked="" type="checkbox"/>
networks.....	<input checked="" type="checkbox"/>	agent-based modeling..	<input checked="" type="checkbox"/>

Although originally published in 1992, this book remains the best general introduction to complexity science.¹ Written in an engaging biographical style, it introduces most of the major topics of complexity science (such as complex adaptive systems, agents, self-organization, emergence, evolution, phase transition, computation, the edge of chaos, and power laws), as well as most of the people involved in its development (including Per Bak, George Cowan, Doyne Farmer, Chris Langton, Murray Gell-Mann, John Holland, Stuart Kauffman, Norman Packard, and Stephen Wolfram).

In it, you will not find a single complexity science equation, graph, picture, or computer code snippet. But from it you will glean something perhaps more important: the frustrations of scientists and economists with traditional scientific methods, the moments of inspiration, the evolution of thought, and the interactions of people that led to Complexity Science.

The sciences of the 21st century

“Complexity, adaptation, upheavals at the edge of chaos – these common themes are so striking that a growing number of scientists are convinced that there is more here than just a series of nice analogies. The movement’s nerve center is a think tank known as the Santa Fe Institute ... The researchers who gather there ... all share the vision of an underlying unity, a common theoretical framework for complexity that would illuminate nature and humankind alike. ... They believe that their application of these ideas is allowing them to understand the spontaneous, self-organizing dynamics of the world in a way that no one ever has before – with the potential for immense impact on the conduct of economics, business, and even politics. They believe that they are forging the first rigorous alternative to the kind of linear, reductionist thinking that has dominated science since the time of Newton – and that has now gone about as far as it can go in addressing the problems of our modern world. They believe they are creating, in the words of Santa Fe Institute founder George Cowan, ‘the sciences of the twenty-first century’.

This is their story.”

from Chapter 1, pp 12-13

¹ For another perspective on the history of complexity science, also with interesting vignettes about its key personalities, you may enjoy reading Lewin (1999). Other readable classics about the development of complexity science are Gleick (2008), Gribbin (2004), Bak (1996), Kelly (1994) and Pagels (1988).

For more technical overviews of complexity science, see Érdi (2008) or Mitchell (2009). For a more advanced overview, see Gros (2008). For a more extensive reference resource, see the 10-volume encyclopedia edited by Meyers (2009). (To find a library near you that carries a copy of the Encyclopedia, search “www.WorldCat.org”.)

2. MANAGING BUSINESS COMPLEXITY

DISCOVERING STRATEGIC SOLUTIONS WITH AGENT-BASED MODELING AND SIMULATION

BY: MICHAEL J. NORTH AND CHARLES M. MACAL (2007)

Structure: 15 chapters, 313 pages, with an index and a bibliography for each chapter

Recommended reading: all chapters

Themes:

complex adaptive systems..	<input checked="" type="checkbox"/>	behavioral economics...	<input type="checkbox"/>
networks.....	<input type="checkbox"/>	agent-based modeling..	<input checked="" type="checkbox"/>

Main ideas:

- Agent-based modeling is a new way to understand complex systems – a way for businesses to view potential futures and to anticipate the likely effects of their decisions on their markets and industries.
- Because of the naturalness of the agent representation and the close similarity of agent models to the predominant paradigm of object-oriented programming, in the future virtually all computer simulations will be agent-based.

This book is the best resource for learning how to do agent-based modeling and simulation.² It is a complete resource, starting from basics for those who know nothing about the subject, and including many examples with sample code on various agent-based modeling platforms. It consists of three sections:

- The first five chapters explain the basics of agent-based modeling.
- Chapters 6 through 10 describe how to construct an agent-based model in different environments.
- Chapters 11 through 15 deal with ancillary issues, such as software development and code verification and validation, that are critical to developing useful agent-based models.

Of particular interest, an extended section of Chapter 5 discusses the differences between agent-based modeling and other modeling types such as systems dynamics, discrete event simulation, participatory simulation, optimization models, statistical modeling, and risk analysis.

For a brief introduction to agent-based modeling, you may enjoy reading the paper *Agent-based modeling and simulation*, by the same authors³

Who, what, where, when, why, and how

“This book addresses the who, what, where, when, why, and how of agents:

- Who should know about agent modeling? Who should do agent modeling? Who should use information generated from agent modeling? ...
- What do agents allow us to do that cannot be done using standard modeling approaches? ...
- Where are the promising applications of ABMS [agent-based modeling and simulation] in the everyday business problems that surround us? ...
- When is it appropriate to use agent-based modeling and simulation? ...
- Why do people use agent-based modeling? ...
- How should one think about agents? How should one go about building agent-based models?

This book provides the answers to these critically important questions for anyone who has heard about agent-based modeling or for those who are considering undertaking an agent-based modeling enterprise.”

from Chapter 1, pp 3-4

² A much shorter resource for learning how to develop agent-based models is Gilbert (2008).

³ Macal & North (2009)

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

3. LINKED

HOW EVERYTHING IS CONNECTED TO EVERYTHING ELSE AND WHAT IT MEANS FOR BUSINESS, SCIENCE, AND EVERYDAY LIFE

BY: ALBERT-LÁSZLO BARABÁSI (2003)

Structure: 16 chapters, 294 pages, with notes (some of which give bibliographical references), and an index

Recommended reading: all chapters

Themes:

complex adaptive systems..	<input type="checkbox"/>	behavioral economics...	<input type="checkbox"/>
networks.....	<input checked="" type="checkbox"/>	agent-based modeling..	<input type="checkbox"/>

Main ideas:

- Many of the complex systems surrounding us are characterized by one universal network architecture – including hubs and the power law distribution – that is simultaneously robust and fragile.
- Many real networks are governed by two laws: growth and preferential attachment. This is why the rich get richer.
- The structure of an organization’s network is responsible for its ability to adapt (or not) to rapidly changing market conditions.
- An understanding of network architecture can lead to new discoveries and strategies for business, politics, and economics.

The importance of networks

“Today we increasingly realize that nothing happens in isolation. Most events and phenomena are connected, caused by, and interacting with a huge number of other pieces of a complex universal puzzle. We have come to see that we live in a small world, where everything is linked to everything else. We are witnessing a revolution in the making as scientists from all different disciplines discover that complexity has a strict architecture. We have come to grasp the importance of networks.”

from *Introduction*, pp 6-7

This book tells the story of the progress of network theory from 1783 to 2002, from Euler’s bridges of Königsberg, through “six degrees of separation”, “small worlds”, Pareto’s 80/20 rule, and network hierarchies, to recent advances. This book is not technical, but it lays the groundwork for understanding more technical works about network theory.⁴ In that regard, it is a better introduction to network theory than Gladwell’s *Tipping point* or Strogatz’s *Sync*.

As an introduction to the author as well as other major scientists in the field, you may enjoy watching the documentary about network theory titled *Connected: the power of six degrees*.⁵

⁴ After reading this book, works such as M. E. J. Newman (2010), M.E.J. Newman (2008), M.E.J. Newman (2003), M.E.J. Newman, et al. (2006), Watts (2003), and Buchanan (2002), Barrat, et al. (2008) would be good next steps.

⁵ This documentary, which originated on TV’s Discovery Channel, may be found as a series of five short videos on YouTube.com, the first of which is: Discovery Channel (2008)

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

4. THE STRUCTURE AND DYNAMICS OF NETWORKS

BY: MARK NEWMAN, ALBERT-LÁSZLO BARABÁSI, AND DUNCAN WATTS (2006)

Structure: 6 chapters, 582 pages, with references and an index.

Recommended reading: chapter 1, introductions to all chapters and chapter sections, chapter 6.

Themes:

complex adaptive systems..	<input type="checkbox"/>	behavioral economics...	<input type="checkbox"/>
networks.....	<input checked="" type="checkbox"/>	agent-based modeling..	<input type="checkbox"/>

This book is a collection of 44 key research papers about network science, with excellent introductions to and summaries of the papers.

Chapter 1 gives a brief history of the study of networks, and describes the new science of networks.

Chapter 2 includes key historical papers, including the short story titled “Chain-links” by Karinthy, the first known mention of “small worlds”, one of Stanley Milgram’s papers about the small world effect, and an interesting paper about the Erdős number.

Chapter 3 presents empirical studies of the world wide web, the Internet, metabolic networks, the web of human sexual contacts, and the structure of scientific collaborations.

Chapter 4 includes papers about network models, including random graph models, small-world models, and models of scale-free networks. In this collection are the seminal papers “Collective dynamics of ‘small-world’ networks” by Watts and Strogatz, and “Emergence of scaling in random networks” by Barabási and Albert.

Chapter 5 presents applications of network science, including applications about contagious disease, susceptibility of networks to attack, and network search algorithms.

Chapter 6 concludes the book with speculations about the future of network science. It concludes with the thought, “... if the science of networks is to have an impact in policy, business, and technology, then research that explicitly addresses the relationship between network properties and the behavior of networked systems will need to be pursued.”

By studying networks, there is much to be learned

“Networks such as the Internet, the World Wide Web, and social and biological networks of various kinds have been the subject of intense study in recent years. From physics and computer science to biology and the social sciences, researchers have found that a great variety of systems can be represented as networks, and that there is much to be learned by studying those networks. The study of the web, for instance, has led to the creation of new and powerful web search engines that greatly outperform their predecessors. The study of social networks has led to new insights about the spread of diseases and techniques for controlling them. The study of metabolic networks has taught us about the fundamental building blocks of life and provided new tools for the analysis of the huge volumes of biochemical data that are being produced by gene sequencing, microarray experiments, and other techniques.

In this book we have gathered together a selection of research papers covering what we believe are the most important aspects of this new science. The papers are drawn from a variety of fields, from many different journals, and cover both empirical and theoretical aspects of the study of networks.”

from the Preface

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

5. A NEW KIND OF SCIENCE

BY: STEPHEN WOLFRAM (2002)

Structure: Preface, 12 chapters, 1,197 pages, with extensive notes (349 pages) and index, but no bibliography

Recommended reading: preface, and chapters 1, 2 (pp. 23-39), 3 (51-70, 105-113), 5 (169-183), 6 (223-254, 261-266, 275-280), 7, 8 (429-432), 9 (433-457), 10 (547-597), 11, 12

Themes:

complex adaptive systems..	<input checked="" type="checkbox"/>	behavioral economics..	<input type="checkbox"/>
networks.....	<input type="checkbox"/>	agent-based modeling..	<input checked="" type="checkbox"/>

Main ideas:

- Simple rules can produce behavior of immense complexity.
- Because most systems that are amenable to mathematical analysis are relatively simple, mathematics is not the best tool to analyze complex systems.
- Simple computer experiments to investigate simple questions are an important part of scientific discovery, and offer a new way to view the operations of complex systems.
- For most complex systems, prediction is impossible.
- Complex systems are computationally equivalent.

This book is the best introduction to the world of cellular automata, and to the far-reaching insights about complex systems that they convey.⁶ Although many of the book's provocative assertions have led to storms of protest⁷, its central message is key for understanding complexity science: simple behavior rules can lead to immense complexity, a phenomenon that is best investigated with computer programs.

Although the book is wide-ranging and fascinating (for example, it offers insights into topics as diverse as free will, space-time, and the ultimate rule of the universe) **for this report's purposes, only the recommended readings given above are directly relevant.**

As an introduction to Stephen Wolfram and this book, you may enjoy watching his presentation of *A new kind of science* at the University of California, San Diego.⁸

⁶ Another excellent book solely about cellular automata is Schiff (2008).

⁷ See, for example, the web page "shell.cas.usf.edu/~wclark/ANKOS_reviews.html". It provides links to dozens of book reviews about *A new kind of science* (most of which are negative) as well as to humorous satires about the book and its author.

⁸ Wolfram (2008), a YouTube.com video, lasting about 90 minutes.

A new approach to complexity

"In the existing sciences much of the emphasis over the past century or so has been on breaking systems down to find their underlying parts, then trying to analyze these parts in as much detail as possible. And particularly in physics this approach has been sufficiently successful that the basic components of everyday systems are by now completely known. But just how these components act together to produce even some of the most obvious features of the overall behavior we see has in the past remained an almost complete mystery. Within the framework of the new kind of science that I develop in this book, however, it is finally possible to address such a question."

from Chapter 1, pp 3-4

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

6. PREDICTABLY IRRATIONAL

THE HIDDEN FORCES THAT SHAPE OUR DECISIONS

BY: DAN ARIELY (2008)

Structure: introduction and 13 chapters; 368 pages, with notes, index, and bibliography

Recommended reading: all chapters

Themes:

complex adaptive systems..	<input type="checkbox"/>	behavioral economics...	<input checked="" type="checkbox"/>
networks.....	<input type="checkbox"/>	agent-based modeling..	<input type="checkbox"/>

Main ideas:

- Expectations, emotions, social norms, and other invisible, seemingly illogical forces skew our reasoning abilities.
- Our reliance on standard economic theory to design personal, national, and global policies may be dangerous.

Written by a behavioral economist, this entertaining book demonstrates the explanatory power of behavioral economics, and the necessity of behavioral experiments for understanding the often irrational – but generally predictable – rules that govern human behavior.

The book’s scope is more general and applicable to the business of actuaries than either of the other popular books *Nudge* (focusing on the choices people make and how their choices can be influenced) or *The mind of the market* (focusing on how behavioral economics affects our daily lives).⁹ The book is an excellent introduction to more technical works on behavioral economics by Kahneman and Tversky.¹⁰

When you read this book it is helpful to pause after each chapter to consider the potential implications of the behavioral economics experiments for actuarial work, and how the behavioral rules might be reflected in agent-based models of actuarial systems.

As an introduction to the author you may enjoy watching his presentation of the book as part of the Authors @ Google program.¹¹

⁹ Thaler & Sunstein (2008) and Shermer (2008)

¹⁰ For example, see the books Gilovich, Griffin, & Kahneman (2002), Kahneman & Tversky (2000), and Kahneman, et al. (1982). You may also be interested to see videos of Kahneman on YouTube.com: Kahneman (2008a) and Kahneman (2008b).

¹¹ Ariely (2008a)

When do headaches vanish?

“Do you know why we so often promise ourselves to diet, only to have the thought vanish when the dessert cart rolls by?

Do you know why we sometimes find ourselves excitedly buying things we don’t really need?

Do you know why we still have a headache after taking a one-cent aspirin, but why that same headache vanishes when the aspirin costs 50 cents?

Do you know why people who have been asked to recall the Ten Commandments tend to be more honest (at least immediately afterward) than those who haven’t? Or why honor codes actually do reduce dishonesty in the workplace?

By the end of this book, you’ll know the answers to these and many other questions that have implications for your personal life, for your business life, and for the way you look at the world. Understanding the answer to the questions about aspirin, for example, has implications not only for your choice of drugs, but for one of the biggest issues facing our society: the cost and effectiveness of health insurance.”

from the Introduction, pp xxi-xxii

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

7. COMPLEX ADAPTIVE SYSTEMS

AN INTRODUCTION TO COMPUTATIONAL MODELS OF SOCIAL LIFE

BY: JOHN H. MILLER AND SCOTT E. PAGE (2007)

Structure: Preface, 12 chapters, Epilogue, two appendices, 263 pages, with an index and bibliography

Recommended reading: all content

Themes:

complex adaptive systems..	<input checked="" type="checkbox"/>	behavioral economics...	<input type="checkbox"/>
networks.....	<input type="checkbox"/>	agent-based modeling..	<input checked="" type="checkbox"/>

Main ideas:

- The science of complexity and its ability to explore the in-between areas (in between the extremes of traditional economics, in between stasis and chaos, control and anarchy, and the continuous and the discrete) is especially relevant for some of the most pressing issues of our modern world.
- While complex systems can be fragile, they also exhibit unusual robustness to changes in their component parts.
- The innate features of many social systems produce complexity.
- The commonly-mentioned “edge of chaos” is not one edge, but rather a multitude of edges contained within a set of attractor rules.

This book provides an excellent introduction to many key models of Complexity Science. It starts with the “standing ovation” model, and progresses through “attack of the killer bees”, the Tiebout world, Conway’s Game of Life, the forest fire model, social cellular automata, Schelling’s segregation model, Schelling’s beach model, Bak’s sandpile model, the Prisoner’s Dilemma, and others. Along the way, the authors discuss related topics such as the pros and cons of computational modeling, what agents are, a basic framework for computational models, the edge of chaos, network structure, self-organized criticality, power laws, fat tails, and game theory.

The two appendices are especially interesting and useful. Appendix A discusses 19 unanswered questions that complexity science must address, such as:

- What does it take for a system to exhibit complexity?
- What makes a system robust?

Appendix B gives best practices for successful computational modeling.

A new approach to complexity

“[The topic of] complex systems has become both a darling of the popular press and a rapidly advancing scientific field. Unfortunately, this creates a gap between popular accounts that rely on amorphous metaphors and cutting-edge research that requires a technical background. Here we hope to provide a point of entry that lies between metaphor and technicalities. Our work focuses on simple examples that are accessible, yet also contain much deeper foundational insights.”

from Chapter 1, p 6

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

8. GROWING ARTIFICIAL SOCIETIES

SOCIAL SCIENCE FROM THE BOTTOM UP

BY: JOSHUA M. EPSTEIN AND ROBERT AXTELL (1996)

Structure: 6 chapters, 208 pages, with appendices, index, and bibliography

Recommended reading: all chapters and appendices

Themes:

complex adaptive systems..	<input checked="" type="checkbox"/>	behavioral economics...	<input type="checkbox"/>
networks.....	<input type="checkbox"/>	agent-based modeling..	<input checked="" type="checkbox"/>

Main ideas:

- Agent-based modeling is the basis of a new kind of generative social science, one that evolves the characteristics of society from the bottom up, based on primitive agents.
- Generative social science provides new insights into the workings of complex social systems. For example, the book’s “Sugarscape” model shows that two basic tenets of traditional economic theory, equilibrium and Pareto optimality, are not realistic.
- All spheres of social life, including demography, economics, cultural change, conflict, and public health, emerge naturally, without top-down specification, from the purely local interactions of individual agents in society.

One of the most cited works in complexity science, this groundbreaking and enduring book demonstrates the power of agent-based modeling, and will generously reward your study.¹² In it, the authors show how an economy can be grown from scratch in a computer world called “Sugarscape”, starting with nothing more than 250 agents, a few basic behavior rules, and an environment with some natural resources. Along the way, it shows that many of the assumptions of macro- and micro-economics are unnecessary or wrong. The Sugarscape model demonstrates: the distribution of wealth, social networks, migration, combat, trade, and disease transmission.

To read about the flowering of generative social science introduced in this book, see the book *Generative Social Science*, also by Joshua Epstein, and also one of the *Top ten Complexity Science books*. In addition to this book, there are other related books that you may enjoy.¹³

Can you grow it?

“The broad aim of this research is to begin the development of a more unified social science, one that embeds evolutionary processes in a computational environment that simulates demographics, the transmission of culture, conflict, economics, disease, the emergence of groups, and agent coadaptation with an environment, all from the bottom up. Artificial society-type models may change the way we think about explanation in the social sciences.

What constitutes an explanation of an observed social phenomenon? Perhaps one day people will interpret the question, “Can you explain it?” as asking “Can you *grow* it?” Artificial society modeling allows us to “grow” social structures *in silico* demonstrating that certain sets of microspecifications are *sufficient to generate* the macrophenomena of interest. And that, after all, is a central aim. As social scientists, we are presented with “already emerged” collective phenomena, and we seek microrules that can generate them. We can, of course, use statistics to test the match between the true, observed structures and the ones we grow. But the ability to grow them – greatly facilitated by modern object-oriented programming – is what is new. Indeed, it holds out the prospect of a new, *generative*, kind of social science.”

from Chapter 1, pp 19-20

¹² It was the most-cited reference in the ten-year history of the Journal of Artificial Societies and Social Simulation. See Meyer, Lorscheid, & Troitzsch (2009).

¹³ See Casti (1997), Axelrod (1997), Holland (1995), and Schelling (2006)

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

9. GENERATIVE SOCIAL SCIENCE

STUDIES IN AGENT-BASED COMPUTATIONAL MODELING

BY: JOSHUA M. EPSTEIN (2006)

Structure: Introduction, nine chapters (each with a “prelude”), and an ending “Coda”; 356 pages, with an index, bibliographies at the end of each chapter, and a CD.

Recommended reading: Introduction, Chapters 1,2,7,9,12,13, and the “Coda”

Themes:

complex adaptive systems..	<input checked="" type="checkbox"/>	behavioral economics...	<input type="checkbox"/>
networks.....	<input type="checkbox"/>	agent-based modeling..	<input checked="" type="checkbox"/>

Main ideas:

- Agent-based computational models are a new scientific instrument that permits a distinctively useful approach to social science.
- Generative explanations in social science are scientifically valid.
- Well-designed generative multi-agent models are more realistic explanations of social phenomena than the “proofs” of economic equilibria or Nash (game theoretic) equilibria.

This book presents some of the concrete achievements of the generative approach to social science that the author’s previous book *Growing artificial societies* (also one of the *Top ten Complexity Science books*) began. For example, Chapter 7 shows how agent-based modeling was used to explain the surprising emergent patterns of retirement after the U.S. Social Security retirement age was changed, and Chapter 12 introduces a model for a smallpox containment strategy.

On the book’s accompanying CD are animations and applets of the models it describes. Computer source code for the models is available from the author.

As an introduction to the author, and a discussion of the smallpox containment model in the book, you may enjoy watching his interview titled *A conversation with Josh Epstein (part 1)*.¹⁴

**If you didn’t grow it,
you didn’t explain its emergence**

“Agent-based models provide computational demonstrations that a given microspecification is in fact sufficient to generate a macrostructure of interest. Agent-based modelers may use statistics to gauge the generative sufficiency of a given microspecification – to test the agreement between real-world and generated macro structures. ... A good fit demonstrates that the target macrostructure – the explanandum – be it wealth distribution, segregation pattern, price equilibrium, norm, or some other macrostructure, is effectively attainable under repeated application of agent-interaction rules: It is effectively computable by agent society. ... Indeed, this demonstration is taken as a necessary condition for explanation itself. To the generativist – concerned with formation dynamics – it does not suffice to establish that, if deposited in some macroconfiguration, the system will stay there. Rather, the generativist wants an account of the configuration’s attainment by a decentralized system of heterogeneous autonomous agents. Thus, the motto of generative social science, if you will, is: If you didn’t grow it, you didn’t explain its emergence ...”

from Chapter 1, p 8

¹⁴ On YouTube.com, the interview is titled *Agent-based modeling and the smallpox example*, J. Epstein (2008a).

Complexity science – an introduction for actuaries

Top ten Complexity Science books continued

10. THE ORIGIN OF WEALTH

EVOLUTION, COMPLEXITY, AND THE RADICAL REMAKING OF ECONOMICS

BY: ERIC D. BEINHOCKER (2006)

Structure: Preface, 18 chapters in four parts, and an Epilogue; 527 pages, with notes, index, and bibliography

Recommended reading: preface, all chapters, and the epilogue

Themes:

complex adaptive systems..	<input checked="" type="checkbox"/>	behavioral economics...	<input checked="" type="checkbox"/>
networks.....	<input checked="" type="checkbox"/>	agent-based modeling..	<input checked="" type="checkbox"/>

Main ideas:

- Economics is in the midst of a revolution that promises to overthrow a century of conventional theory; Complexity Science is the revolution's primary enabler.
- The behavior of economic systems is part of a universal class of evolutionary behavior, with common evolutionary laws and algorithms. Wealth is created through these evolutionary algorithms acting on technologies, social institutions, and businesses. Indeed, evolution is the formula that lies behind all the order, complexity, diversity, and wealth in the economic world.
- The economy was misclassified by traditional economics as an equilibrium system, whereas in fact it is a complex adaptive system.
- A modern economy's behavior cannot be forecasted for more than a very short time.
- At the core of any economic theory must be a realistic theory of human behavior, an ingredient currently missing in traditional economics.
- Much of the volatility we see in current economies (business cycles, growth discontinuities, inflation, etc.) is generated by the dynamics of people's behavior, rather than by exogenous shocks.

This unique tour de force weaves the themes of complex adaptive systems, networks, behavioral economics, and agent-based modeling into a highly readable account of how Complexity Science is transforming the field of economics.

For a brief introduction to the author, see his interview with Richard Dawkins.¹⁵

Complexity Economics

"As I write this, the field of economics is going through its most profound change in over a hundred years. I believe that this change represents a major shift in the intellectual currents of the world that will have a substantial impact on our lives and the lives of generations to come. ...

Despite the importance of economic thinking, few people outside the hushed halls of academia are aware of the fundamental changes under way in the field today. This book is the story of what I will call the Complexity Economics revolution: what it is, what it tells us about the deepest mysteries in economics, and what it means for business and for society at large."

from the Preface, pp xi-xii

¹⁵ See Beinhocker (2009) for part 1 (of 3) of the interview sequence.

ESSENTIAL RESOURCES

A. INTRODUCTION

This section lists the resources that, in my opinion, are the most useful for actuaries to learn more about complexity science and apply it in their work. I found these resources using the process described in the section *Finding the essential resources*. They include important websites, journals, conferences, e-mail lists, books, articles, and videos.

B. WEBSITES

The two most important websites for material about Complexity Science are:

- **www.santafe.edu** The site for the Santa Fe Institute, the premier research institution for Complexity Science. Visit this site to keep abreast of new developments in the field.
- **www.econ.iastate.edu/tesfatsi/ace.htm** Devoted to agent-based computational economics, this site provides introductory material and links to important related sites.

C. JOURNALS

The two most prominent journals focusing on Complexity Science are:

- **Journal of Artificial Societies and Social Simulation (JASSS)** A free journal, available only on the Internet. You can sign up to receive an e-mail when each issue is published.
- **Computational and Mathematical Organization Theory (CMOT)** A subscription-based journal.

D. CONFERENCES

The two most important conferences related to social simulation and agent-based modeling are:

- **The Winter Simulation Conference (WSC)**
- **The World Congress on Social Simulation** This biannual (even years) congress is organized by three regional societies:
 - European Social Simulation Association (ESSA)
 - North American Association for Computational Social and Organization Sciences (NAACSOS), recently reorganized as the Computational Social Science Society (CSSS)
 - Pacific Asian Association for Agent-Based Approach in Social Systems Sciences (PAAA)

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

E. E-MAIL LISTS

The most important e-mail list is the SIMSOC e-mail distribution list. This list will notify you about forthcoming conferences and workshops related to agent-based modeling. To subscribe to the e-mail list, or to view the archives, visit the website “www.jiscmail.ac.uk/lists/simsoc.html”.

F. BOOKS, ARTICLES, AND VIDEOS

Following are the essential introductory books, articles, and videos about complexity science for actuaries. The list is sorted alphabetically by the primary author’s last name. As a guide to each reference, its last lines provide the following information:

- The area of actuarial interest, such as Health, Pensions, Risk Management, Insurance, and General (for references of general interest).
- A brief description of the reference’s contents.
- The type of reference, such as Book, Book Section, Article, Conference Paper, Report, and Video.

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

Aaron, H. J. (1999). *Behavioral dimensions of retirement economics*. Washington D.C. New York: Brookings Institution Press; Russell Sage Foundation, 289 pages.

(Pensions: Presents a collection of reports about the application of behavioral economics to retirement issues - Book)

Adamic, L. A. (2002). *Zipf, power-laws, and Pareto - a ranking tutorial*: HP Labs - Information Dynamics Lab, 5 pages.

(General: Shows that the power law, Zipf's law, and the Pareto distribution are mathematically equivalent, and discusses when to use each - Report)

Aldrich, C. (2009). *The complete guide to simulations & serious games*: John Wiley & Sons, Inc.,

(General: Provides an overview of serious games and how they are used for education - Book)

Alkemper, J., & Mango, D. F. (2005). Concurrent simulation to explain reinsurance market price dynamics. *Risk Management, November 2005* (6), 13-17.

(Insurance: Reports on an agent-based model the authors developed to provide insights into the dynamics of the property-casualty reinsurance market - Article)

Ariely, D. (2008a). Authors @ Google: Dan Ariely. Retrieved from www.youtube.com. 56 minutes.

(General: Presents Dan Ariely's discussion of his book - Video)

Ariely, D. (2008b). *Predictably irrational: the hidden forces that shape our decisions* (1st ed.). New York, NY: Harper, 280 pages.

(General: Offers insights from behavioral economics into the patterns that cause people to make the same mistakes repeatedly - Book)

Ariely, D. (2008 - 2009). Predictably irrational series. Retrieved from

www.youtube.com/view_play_list?p=B2424C40DE1C0D14&search_query=predictably+irrational.

1 - 5 minutes each.

(General: Presents Dan Ariely's discussion of his book and behavioral economics in general - Video series)

Axelrod, R. M. (1997). *The complexity of cooperation: agent-based models of competition and collaboration*. Princeton, N.J.: Princeton University Press, 232 pages.

(General: Presents a collection of Axelrod's essays primarily about game theory in complexity science - Book)

Bak, P. (1996). *How nature works: the science of self-organized criticality*. New York, NY, USA:

Copernicus, 212 pages.

(General: Introduces the concept of self-organized criticality - Book)

Barabási, A.-L. (2003). *Linked: how everything is connected to everything else and what it means for business, science, and everyday life*. New York: Plume, 294 pages.

(General: Introduces network theory - Book)

Barabasi, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286, 509-512.

(General: Seminal article that established preferential attachment and power law distributions as fundamental properties of many real-world networks - Article)

Barrat, A., Barthelemy, M., & Vespignani, A. (2008). *Dynamical processes on complex networks*.

Cambridge, UK ; New York: Cambridge University Press, 347 pages.

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

(General: Explains the effects of complex network patterns on dynamical phenomena - Book)

Basole, R. C., & Rouse, W. B. (2008). Complexity of service value networks. *IBM Systems Journal*, 47 (1), 53-70.

(Health: Compares the complexity of various economic sectors to the healthcare sector - Article)

Bass, T. A. (1999). *The predictors* (1st ed.). New York: H. Holt and Co., viii, 309 p.

(General: Tells the story of The Prediction Company, from its inception - Book)

Beinhocker, E. D. (2006). *The origin of wealth: evolution, complexity, and the radical remaking of economics*. Boston, Mass.: Harvard Business School Press, 527 pages.

(General: Shows how complexity science is transforming the field of economics - Book)

Beinhocker, E. D. (2009). Interview with Richard Dawkins: The Genius of Darwin (part 1 of 3).

Retrieved from www.youtube.com. 8 minutes.

(General: Presents Eric Beinhocker's interview with Richard Dawkins - Video)

Bonabeau, E. (2002). Predicting the unpredictable. *Harvard Business Review*, 2002 (March), 5-11.

(General: Discusses how to use agent-based modeling when prediction is impossible - Article)

Boucek, C. H., & Conway, T. P. (2003). *Dynamic pricing analysis*. Paper presented at the Casualty Actuarial Society Forum - Winter 2003, 18 pages.

(Insurance: Describes an agent-based simulation model that the authors used to estimate the impact that a pricing rate change will have on a company's policyholder retention and resulting profitability - Conference paper)

Buchanan, M. (2002). *Nexus: small worlds and the groundbreaking science of networks* (1st ed.). New York: W.W. Norton, 235 pages.

(General: Introduces network theory - Book)

Buchanan, M. (2009). Meltdown modelling: could agent-based computer models prevent another financial crisis? *Nature*, 460 (August 6, 2009), 680-682.

(General: Makes the case that Complexity Science could prevent another financial crisis - Article)

Carey, J. (2006). Medical guesswork. *BusinessWeek* (May 29, 2006).

(Health: Describes how the Archimedes model is helping to take the guesswork out of medicine - Journal article)

Casti, J. L. (1997). *Would-be worlds: how simulation is changing the frontiers of science*. New York: J. Wiley, 242 pages.

(General: Discusses how simulation of "artificial worlds" is changing the nature of scientific work - Book)

Clarke, A. C. (2008). Fractals - the colors of infinity [Video]. Retrieved from www.youtube.com. 53 minutes.

(General: Arthur C. Clarke introduces the world of fractals, with support from Benoit Mandelbrot, Stephen Hawking, and others. - Video)

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2007). *Power-law distributions in empirical data*: Santa Fe Institute, 43 pages.

(General: Shows how to correctly analyze data for power law behavior - Report)

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

Conway, J. (2007a). John Conway talks about the Game of Life (part 1 of 2) [Video]. Retrieved from www.YouTube.com. 4 minutes.

(General: Presents John Conway talking about his Game of Life - Video)

Conway, J. (2007b). John Conway talks about the Game of Life (part 2 of 2) [Video]. Retrieved from www.YouTube.com. 2 minutes.

(General: Presents John Conway talking about his Game of Life - Video)

Crabb, R. R., & Shapiro, A. F. (1996). Managing the insurance enterprise: an interactive computer game. *Actuarial Research Clearing House*, 1 (1996), 279-289.

(Insurance: Describes a serious game to train students how to manage an insurance company - Article)

Danis, M. (2003). *The CHAT project: choosing healthplans all together*: National Institutes of Health,

(Health: Describes the CHAT project - Article)

Diamond, P. A., Vartiainen, H., & Yrjö Jahnssonin säätiö. (2007). *Behavioral economics and its applications*. Princeton, N.J.: Princeton University Press, xvi, 312 p.

(General: Describes many applications of behavioral economics, including applications in health care, finance, economics, and welfare policy - Book)

Discovery Channel. (2008). Connected: the power of six degrees (on YouTube.com as *A documentary on networks, social and otherwise*) (part 1 of 5): Discovery Channel. Retrieved from www.youtube.com. 9 minutes.

(General: Introduces network theory - Video)

Eddy, D., & Schlessinger, L. (2003). Archimedes - a trial-validated model of diabetes. *Diabetes Care*, 26 (11 - November 2003), 3093-3101.

(Health: Provides an overview of the model Archimedes - Journal article)

Epstein, J. (2008a). Agent-based modeling and the smallpox example. Retrieved from www.youtube.com. 6 minutes.

(Health: Presents Josh Epstein discussing his agent-based smallpox model - Video)

Epstein, J. (2008b). How organizations adapt to new environments. Retrieved from www.youtube.com. 4 minutes.

(General: Presents Josh Epstein discussing his agent-based model of the best organizational structures for adapting to environmental change. - Video)

Epstein, J. M. (2006). *Generative social science: studies in agent-based computational modeling*. Princeton: Princeton University Press, 356 pages.

(General: Presents a collection of reports that are examples of generative social science - Book)

Epstein, J. M. (2008). Why model? *Journal of artificial societies and social simulation*, 11 (412), 5 pages.

(General: Addresses enduring misconceptions about modeling and gives sixteen reasons other than prediction to build models - Article)

Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies: social science from the bottom up*. Washington, D.C.: Brookings Institution Press, 208 pages.

(General: Demonstrates the power of agent-based modeling to help us understand social and economic systems - Book)

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

Érdi, P. (2008). *Complexity explained*. Berlin: Springer, 397 pages.

(General: Introduces complex adaptive systems, mainly from a mathematical perspective - Book)

Fausett, L. V. (1994). *Fundamentals of neural networks : architectures, algorithms, and applications*.

Englewood Cliffs, NJ: Prentice-Hall, xvi, 461 p.

(General: Provides a non-technical introduction to neural networks - Book)

Gardner, M. (1970). The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 1970 (October).

(General: Describes, for the first time, The Game of Life - Article)

Gilbert, G. N. (2008). *Agent-based models*. Los Angeles: Sage Publications, 98 pages.

(General: Provides a brief practical introduction to agent-based modeling - Book)

Gilovich, T., Griffin, D. W., & Kahneman, D. (2002). *Heuristics and biases: the psychology of intuitive judgement*. Cambridge, U.K.; New York: Cambridge University Press, 857 pages.

(General: Provides a collection of research articles about behavioral economics - Book)

Gionta, G. (2000). *Insurance World 2 - A complex model to manage risk in the age of globalization*,

(General: Provides a detailed description of the Insurance World model - Report)

Gladwell, M. (2002). *The tipping point: how little things can make a big difference* (1st Back Bay pbk. ed.).

Boston: Back Bay Books, 301 pages.

(General: Demonstrates the non-linearity of social systems by giving examples of how small changes can have big effects - Book)

Gleick, J. (2008). *Chaos: making a new science* (20th anniversary ed.). New York, N.Y.: Penguin Books, 360 pages.

(General: Traces the development of chaos theory - Book)

Gribbin, J. R. (2004). *Deep simplicity: bringing order to chaos and complexity* (1st U.S. ed.). New York: Random House, 275 pages.

(General: Introduces complex adaptive systems - Book)

Gros, C. (2008). *Complex and adaptive dynamical systems: a primer* (1st ed.). New York: Springer, 262 pages.

(General: Introduces more advanced topics in complex adaptive systems, from a mathematical perspective - Book)

Haldane, A. G. (2009). *Rethinking the financial network*, 41 pages.

(General: Applies insights from complexity science to the financial sphere - Report)

Heath, B., Hill, R., & Ciarallo, F. (2008). A survey of agent-based modeling practices (January 1998 to July 2008). *Journal of artificial societies and social simulation*, 12 (8), 42 pages.

(General: Surveys agent-based modeling practices - Article)

Hickman, J. C. (1997). Introduction to actuarial modeling. *North American Actuarial Journal*, 1 (3), 1-5.

(General: Describes the results of the conference titled "Actuarial and financial modeling: toward a new science" held in 1996 - Article)

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

Holland, J. H. (1995). *Hidden order: how adaptation builds complexity*. Reading, Mass.: Addison-Wesley, 185 pages.

(General: Introduces complex adaptive systems, with an emphasis on the concept of adaptation - Book)

Holland, J. H. (2008). Modeling complex adaptive systems: Case Western Reserve University. Retrieved from www.youtube.com. 71 minutes.

(General: Presents John Holland talking about how to model complex adaptive systems - Video)

Horgan, J. (1996). *The end of science : facing the limits of knowledge in the twilight of the scientific age*. Reading, Mass.: Addison-Wesley Pub., x, 308 p.

(General: Makes the case that all the important scientific discoveries have been made, and that Complexity Science in particular has nothing important to add - Book)

Johansen, A. (1996). A simple model of recurrent epidemics. *Journal of theoretical biology*, 178, 45-51.

(Health: Describes a cellular automaton model of disease spreading - Article)

Kahn, J. (2009). Modeling human drug trials - without the human. *Wired Magazine*, 404 (December 2009).

(Health: Describes the Archimedes model, especially how it is used to simulate drug trials - Journal article)

Kahn, R., Alperin, P., Eddy, D., Borch-Johnsen, K., Buse, J., Feigelman, J., et al. (2010). Age at initiation and frequency of screening to detect type 2 diabetes: a cost-effectiveness analysis. *The Lancet*, 375 (April 17, 2010), 1365-1374.

(Health: Describes the Archimedes simulation of various type 2 diabetes screening strategies - Journal article)

Kahneman, D. (2008a). Explorations of the mind - Intuition: the marvels and the flaws, *Hitchcock Lectures*. Retrieved from www.youtube.com. 56 minutes.

(General: Presents Daniel Kahneman talking about behavioral economics - Video)

Kahneman, D. (2008b). Explorations of the mind - Well-being, *Hitchcock Lectures*. Retrieved from www.youtube.com. 59 minutes.

(General: Presents Daniel Kahneman talking about behavioral economics - Video)

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge ; New York: Cambridge University Press, 555 pages.

(General: Provides a collection of research articles about behavioral economics - Book)

Kahneman, D., & Tversky, A. (2000). *Choices, values, and frames*. New York; Cambridge, UK: Russell sage Foundation; Cambridge University Press, 840 pages.

(General: Provides a collection of research articles about behavioral economics - Book)

Kauffman, S. A. (1995). *At home in the universe: the search for laws of self-organization and complexity*. New York: Oxford University Press, 321 pages.

(General: Introduces complex adaptive systems - Book)

Kelly, K. (1994). *Out of control: the rise of neo-biological civilization*. Reading, MA: Addison-Wesley, 521 pages.

(General: Introduces complex adaptive systems - Book)

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

Kornhauser, D., Wilensky, U., & Rand, W. (2009). Design guidelines for agent based model visualization. *Journal of artificial societies and social simulation*, 12 (2), 27 pages.

(General: Shows how to use visualization for analyzing and presenting agent-based modeling results - Article)

Law, A. M. (2009). *How to build valid and credible simulation models*. Paper presented at the 2009 Winter Simulation Conference, 10 pages.

(General: Presents techniques for building valid and credible simulation models. - Conference paper)

Lewin, R. (1999). *Complexity: life at the edge of chaos* (2nd ed.). Chicago, Ill.: University of Chicago Press, 234 pages.

(General: Introduces complex adaptive systems - Book)

Lewis, A. A. (1985). On effectively computable realizations of choice functions. *Mathematical Social Sciences* (10), 43-80.

(General: Proves that it is not possible to behave with full rationality - Article)

Macal, C. M., & North, M. J. (2009). *Agent-based modeling and simulation*. Paper presented at the 2009 Winter Simulation Conference, 13 pages.

(General: Introduces agent-based modeling and simulation - Conference paper)

Mandelbrot, B. B. (1983). *The fractal geometry of nature* (Updated and augm. ed.). New York: W.H. Freeman, 468 p.

(General: A beautiful book that describes fractals and shows where they appear in nature - Book)

Mango, D. F. (2004). The agents are coming. *The Actuarial Review*, 31 (1), 23-24.

(General: Argues that agent-based simulation is an important new method for actuaries to test the impact of changes in rates, plans, laws, sales compensation, and regulatory policy - Article)

Mango, D. F. (2005). Risk management research imperatives. *North American Actuarial Journal* (July 2005), iii-v.

(Risk management: Calls for new research in risk management - Article)

Meyer, M., Lorscheid, I., & Troitzsch, K. G. (2009). The development of social simulation as reflected in the first ten years of JASS: a citation and co-citation analysis. *Journal of artificial societies and social simulation*, 12 (4), 20 pages.

(General: Reviews the development of social simulation literature - Article)

Meyers, R. A. (Ed.) (2009) *Encyclopedia of complexity and systems science* (1st ed.). New York: Springer.

(General: Presents articles about most areas of complexity science. - Encyclopedia)

Microinsurance. (2008). *Microinsurance - CHAT - Choosing healthplans all together*. Retrieved from www.youtube.com. 6 minutes.

(General: Presents an overview of the application of CHAT in rural india - Video)

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: an introduction to computational models of social life*. Princeton, N.J.: Princeton University Press, 263 pages.

(General: Introduces the theoretical foundations of complex adaptive systems theory - Book)

Mitchell, M. (1996). *An introduction to genetic algorithms*. Cambridge, Mass.: MIT Press, 205 pages.

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

(General: Covers the theory and history of genetic algorithms - Book)

Mitchell, M. (2009). *Complexity: a guided tour*. Oxford, U.K.; New York: Oxford University Press, 349 pages.

(General: Provides an overview of complexity science - Book)

Newman, M. E. J. (2003). *The structure and function of complex networks*: University of Michigan, 58 pages.

(General: Introduces network theory and reviews developments in the field - Report)

Newman, M. E. J. (2008). The physics of networks. *Physics Today, November 2008*, 33-38.

(General: Describes approaches to quantify network patterns and to determine what such patterns mean for the functioning of a system that a network represents. - Article)

Newman, M. E. J. (2010). *Networks : an introduction*. Oxford ; New York: Oxford University Press, xi, 772 p.

(General: The new 'bible' of network science, this book coherently presents the different strands of network science and highlights the interconnections among them - Book)

Newman, M. E. J., Barabási, A.-L., & Watts, D. J. (2006). *The structure and dynamics of networks*. Princeton: Princeton University Press, 582 pages.

(General: Delves into network details - Book)

North, M. J., & Macal, C. M. (2007). *Managing business complexity: discovering strategic solutions with agent-based modeling and simulation*. Oxford ; New York: Oxford University Press, 313 pages.

(General: Shows how to develop agent-based models - Book)

NOVA. (2007). Emergence - complexity from simplicity, order from chaos (1 of 2). Retrieved from www.youtube.com. 5 minutes.

(General: Introduces complex adaptive systems - Video)

Nowak, A., & Lewenstein, M. (1996). Modeling social change with cellular automata. In Hegselmann (Ed.), *Modeling & simulation in the social sciences from a philosophical point of view* (Boston: Kluwer.

(General: Describes and gives modeling examples of social impact theory - Book section)

Orrell, D. (2007). *The future of everything : the science of prediction* (1st Thunder's Mouth Press ed.). New York: Thunder's Mouth Press, 449 p.

(General: Compares the difficulties of predicting the weather, health, and wealth - Book)

Pagels, H. R. (1988). *The dreams of reason: the computer and the rise of the sciences of complexity*. New York: Simon and Schuster, 352 p.

(General: Covers the early history and wider implications of complexity science - Book)

Poundstone, W. (1985). *The recursive universe : cosmic complexity and the limits of scientific knowledge*. Chicago: Contemporary Books, 252 p.

(General: Gives an in-depth analysis of the Game of Life - Book)

Purdy, J. A. (2007). Getting serious about digital games in learning. *Corporate University Journal* (1), 3-6.

(General: Gives an overview of the uses of serious games in business and education - Article)

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

Rauch, J. (2002). Seeing around corners. *The Atlantic Monthly* (April 2002), 35-48.

(General: An introduction to the applications of agent-based modeling - Article)

Regis, E. (2003). *The info mesa : science, business, and new age alchemy on the Santa Fe Plateau* (1st ed.).

New York: W.W. Norton, 268.

(General: Describes the many businesses that have spun off from the Santa Fe Institute - Book)

Rhodes, C. J., Jensen, H. J., & Anderson, R. M. (1997). On the critical behaviour of simple epidemics. *Proceedings of the Royal Society London*, 264, 1639-1646.

(Health: Describes the power-law characteristics of epidemics for an island population - Article)

Sanders, I. T., & McCage, J. A. (2003). *The use of Complexity Science: a survey of Federal departments and agencies, private foundations, universities, and independent education and research centers*: Washington Center for Complexity & Public Policy, 62 pages.

(General: Surveys the organizations involved in Complexity Science and how the new science is being used - Report)

Sargent, R. G. (2009). *Verification and validation of simulation models*. Paper presented at the 2009 Winter Simulation Conference, 15 pages.

(General: Discusses various approaches to verify and validate simulation models - Conference paper)

Schelling, T. C. (2006). *Micromotives and macrobehavior*. New York: Norton, 270 pages.

(General: Shows how seemingly trivial individual behaviors can lead to important aggregate results - Book)

Schiff, J. L. (2008). *Cellular automata: a discrete view of the world*. Hoboken, N.J.: Wiley-Interscience, 252 pages.

(General: Introduces cellular automata and their application - Book)

Schlessinger, L., & Eddy, D. (2001). Archimedes: a new model for simulating health care systems - the mathematical formulation. *Journal of Biomedical Informatics*, 35 (2002), 37-50.

(Health: Describes the mathematical formulation to model human biology, disease, and healthcare interventions in Archimedes - Journal article)

Segre-Tossani, L. (2003). Simulation technology for managing risk. *Risks and Rewards Newsletter*, February 2003 (41), 4-10.

(Risk management, Insurance: Describes the agent-based model InsuranceWorld - Article)

Serious Games Institute. (2007). The Serious Games Institute: Building infrastructure for the serious games sector. Retrieved from www.youtube.com. 3 minutes.

(General: Presents an overview of the Serious Games Institute - Video)

Shermer, M. (2008). *The mind of the market: compassionate apes, competitive humans, and other tales from evolutionary economics* (1st ed.). New York: Times Books, 308 pages.

(General: Explains human economic behavior using behavioral economics research results - Book)

Shumrak, H. M., & Darley, V. (1999). *Applying complex adaptive systems to actuarial problems*. Paper presented at the 1999 Valuation Actuary Symposium Proceedings, 28 pages.

(General: Describes an agent-based simulation to model customer lapse behavior - Conference paper)

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

Shumrak, H. M., Greenbaum, M., Darley, V., & Axtell, R. (1999). *Modeling annuity policyholder behavior using behavioral economics and complexity science*, 12 pages.

(Insurance: Describes a model of policyholder behavior that combines behavioral economics and agent-based modeling - Report)

Smith, L. M., & Segre-Tossani, L. (2003). *Applications of advanced science in the new era of risk modeling*. Paper presented at the 2003 Thomas P. Bowles, Jr. Symposium - April 10-11,

(Risk management: Presents the case that agent-based modeling is better than traditional actuarial methods to deliver credible analyses in a risk environment characterized by multiple correlations, extreme events, and cascading risks - Conference paper)

Sobkowicz, P. (2003). Opinion formation in networked societies with strong leaders. *Complexity Digest*, 2003 (48).

(General: Provides details about social impact theory - Article)

Strip, D., Backus, G., Strickland, J., & Schoenwald, D. (2005). *Modeling the US healthcare system: predicting the consequences of policy decisions through computational models*: Sandia Corporation, 4 pages.

(Health: Argues for creating an agent-based model of the entire U.S. healthcare system in order to better understand it - Report)

Studnicki, J., Eichelberger, C., & Fisher, J. (2009). Complex adaptive systems: how informed patient choice influences the distribution of complex surgical procedures. In Z. W. Ras & W. Ribarsky (Eds.), *Advances in information and intelligent systems - studies in computational intelligence* (Vol. 251): Springer. 17 pages.

(Health: Describes a model the authors developed to study how informed patient choices can influence the distribution of surgical volume for complex procedures - Book section)

Takadama, K., & Shimohara, K. (2002). The hare and the tortoise: cumulative progress in agent-based simulation. In A. Namatame, T. Terano & K. Kurumatani (Eds.), *Agent-based approaches in economic and social complex systems* (Amsterdam: IOS Press). pages 3-14.

(General: Argues for the slow, careful, and thorough development of agent-based modeling and simulation - Book section)

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: improving decisions about health, wealth, and happiness*. New Haven: Yale University Press, 293 pages.

(General: Shows how human behavior can be molded using behavioral economics research results - Book)

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, Conn.: Graphics Press, 197 p.

(General: Provides guidance about displaying quantitative information - Book)

Tufte, E. R. (2006a). *Beautiful evidence*. Cheshire, Conn.: Graphics Press, 213 p.

(General: Provides guidance about displaying quantitative information - Book)

Tufte, E. R. (2006b). *The cognitive style of PowerPoint: pitching out corrupts within* (2nd ed.). Cheshire, Conn.: Graphics Press, 31 p.

(General: Makes the case that PowerPoint is misused, and shows better ways to make a presentation - Booklet)

Waldrop, M. M. (1992). *Complexity: the emerging science at the edge of order and chaos*. New York: Simon & Schuster, 380 pages.

(General: Introduces complexity science - Book)

Complexity science – an introduction (and invitation) for actuaries

Essential resources continued

Wang, S., & Mango, D. F. (2003). Research outside the actuarial comfort zone. *Actuarial Review*, 30 (1 (February 2003)).

(General: Argues that actuaries should take a leadership role in addressing systemic societal problems - Article)

Watts, D. J. (2003). *Six degrees: the science of a connected age* (1st ed.). New York: Norton, 368 pages.

(General: Introduces network theory - Book)

Wei, Y.-m., Ying, S.-j., Fan, Y., & Wang, B.-H. (2003). The cellular automaton model of investment behavior in the stock market. *Physica*, 325 (2003), 507-516.

(General: Applies an automaton model to study stock market investment behavior - Article)

Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media, 1197 pages.

(General: Introduces cellular automata and their application - Book)

Wolfram, S. (2008). A new kind of science - Stephen Wolfram [Video]: University of California, San Diego. Retrieved from www.youtube.com. 86 minutes.

(General: Presents Stephen Wolfram talking about his book - Video)

Wragg, T. (2006). *Modelling the effects of information campaigns using agent-based simulation*: Australian Government Department of Defence - Technology Organisation, 58.

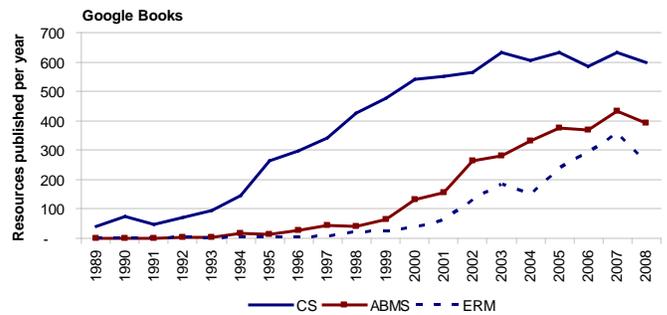
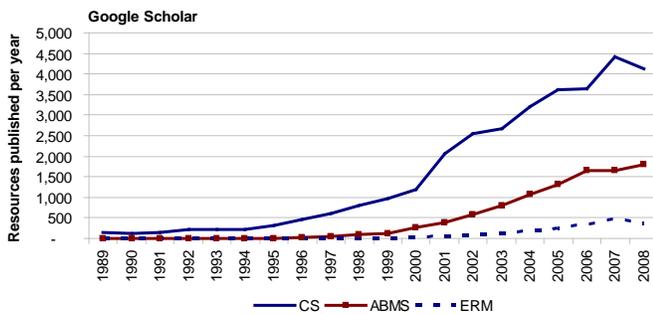
(General: Demonstrates an agent-based model of an information campaign regarding vaccination in India - Report)

FINDING THE ESSENTIAL RESOURCES

A. INTRODUCTION

Even though Complexity Science is young, its literature is vast and scattered. This section describes the process I followed during December 2009 and January 2010 to search through this literature and find the essential resources for actuaries.

As a measure of the literature’s size, on Google Web one finds over 500,000 resources related to complexity science (CS) and agent-based modeling and simulation (ABMS). Also, according to Google Scholar and Google Books, during the twenty-year period 1989 – 2008, more than 50,000 related articles and books were published (see the charts below). By contrast, during the same twenty-year period, only about 4,000 articles and books were published about another young field of interest to actuaries, Enterprise Risk Management (ERM).¹



Not only are the Complexity Science resources numerous, they are also widely dispersed. Mirroring the variety of fields that Complexity Science touches, its resources are scattered among numerous journals, websites, conferences, publishers, and institutions, and are classified under many domains. For example, a recent survey of agent-based modeling applications found 279 articles from 92 publication outlets classified under 9 domains (economics, social science, business, etc.).²

¹ These results were obtained by searching for the terms “complexity science”, “complex adaptive system(s)”, “agent based model(ing)”, and “enterprise risk management” on Google Web, Google Scholar, and Google Books.

² Heath, Hill, & Ciarallo (2008)

B. SEARCH GOALS

The primary goal of this literature search is to compile resources to help actuaries apply the tools and insights of Complexity Science to their work.

A secondary goal is to demonstrate how an actuary can perform a structured literature search, a powerful aid that actuaries often neglect. Accordingly, this literature search includes only databases and search tools that are readily available to actuaries; it includes no purely academic or subscriber databases (such as the ISI Web of Knowledge).

C. SEARCH PROCEDURE

The search process consists of six sequential steps:

- 1. Develop search terms** The first step was to develop search terms. “Search terms” are the words or phrases used to find resources in a database (see sidebar). The process to develop search terms is largely trial and error: one tries a variety of potential terms to find those that yield the most useful and manageable results. As examples, for this search, the term “complexity” yielded too many resources to sort through, while the term “complexity theory” yielded too few. The terms used for this search are listed in Section D – *Search terms*.
- 2. Identify relevant databases** The next step was to identify databases that would provide useful resources, and that actuaries can easily access. After identifying obvious choices such as online databases of professional actuarial organizations and Google Web, finding other databases was largely a matter of trial and error. For example, almost by accident I discovered many valuable resources through Google Video. The databases used for this search are listed in Section E – *Databases*.
- 3. Find primary resources** Using the search terms and databases developed in the first two steps, the next step was to search the databases for potentially useful resources. For example, I searched the Society of Actuaries online database for all resources containing the search term “complexity science”. All the resources returned by a database are called “hits” (see sidebar).

Search terminology

“**Search terms**” are words or phrases related to a topic that are used to find resources on that topic within a database. For example, in many databases I used the search term “complexity science” to find resources related to complex adaptive systems.

“**Hits**” are the resources found in a database from using search terms.

“**Primary resources**” are those resources of potential interest to actuaries that I found directly from searching a database. For example, I found the article “Applying complex adaptive systems to actuarial problems” directly from searching the Society of Actuaries online database. Thus, it is a primary resource.

“**Referenced resources**” are those resources potentially of interest to actuaries that are mentioned within a primary resource. For example, within the article “Applying complex adaptive systems to actuarial problems”, the authors mention an article titled “Modeling annuity policy holder behavior using behavioral economics and complexity science”. This second article is thus a “referenced resource”.

Complexity science – an introduction (and invitation) for actuaries

Finding the essential resources *continued*

C. SEARCH PROCEDURE *CONTINUED*

3. Find primary resources *continued*

By reviewing each hit (or, with Google searches, each of the top 200 hits), I determined its potential usefulness. Potentially useful hits are called “primary resources” (see sidebar on the previous page).

4. Find referenced resources

Next, I reviewed each resource mentioned in the text of every primary resource, to determine its potential usefulness. Such potentially useful resources are called “referenced resources” (see sidebar on the previous page).

For more search details, you can find the complete list of primary and referenced resources, together with a detailed chronicle of the process I used to find them, on the SOA web page for this report.

Section F – *Search results* summarizes the search results.

5. Select the essential resources

The next step was to select from the primary and referenced resources those that are the most important for actuaries to study. I selected these after reading the primary and referenced resources obtained in steps 3 and 4. The essential resources, numbering about one hundred, are listed in the section of this report titled *Essential resources*.

6. Select the top ten books

From among the essential resources, the final step was to select and annotate the ten most important books. These are the most important books, in my opinion, to help actuaries learn more about Complexity Science. They are discussed in the section of this report *Top ten Complexity Science books*.

Complexity science – an introduction (and invitation) for actuaries

Finding the essential resources *continued*

D. SEARCH TERMS

Following are the search terms I used:

<u>Complexity science (CS) search terms</u>	<u>Agent-based modeling and simulation (ABMS) search terms</u>	<u>Actuarial search terms</u>
complexity science complex adaptive system complex adaptive systems network theory network science behavioral economics behavioural economics	agent-based agent based multi-agent multi agent multiagent	actuarial actuaries actuary asset allocation investment insurance reinsurance pension retirement healthcare health care risk management

For actuarial databases, such as the Society of Actuaries online database, I used only CS and ABMS search terms. For other databases, I used the following search term combinations:

- (CS search terms) AND (Actuarial search terms)
- (ABMS search terms) AND (Actuarial search terms)

Search terms were grouped as indicated in the table above (e.g., insurance and reinsurance were grouped together in a search).

The complete list of search term combinations employed is given in Section F – *Search results*.

Complexity science – an introduction (and invitation) for actuaries

Finding the essential resources continued

E. DATABASES

Following are the databases I searched.

<p><u>A. Actuarial online databases</u></p> <ol style="list-style-type: none">1. Society of Actuaries (SOA) (www.soa.org)2. Casualty Actuarial Society (CAS) (www.casact.org)3. International Actuarial Association (IAA) (www.actuaries.org)4. Institute of Actuaries/Faculty of Actuaries (UK) (www.actuaries.org.uk)5. Institute of Actuaries of Australia (www.actuaries.asn.au) <p><u>B. Books</u></p> <ol style="list-style-type: none">1. Library of Congress (www.loc.gov)2. Google Books (www.google.com)3. Santa Fe Institute Book List (www.santafe.edu/research/publications/publications-book-list.php) <p><u>C. Internet</u></p> <ol style="list-style-type: none">1. Google Web (www.google.com)	<p><u>D. Journals</u></p> <ol style="list-style-type: none">1. Google Scholar (www.google.com)2. PubMed www.ncbi.nlm.nih.gov/pubmed3. Santa Fe Institute Working Papers (www.santafe.edu/research/publications/wplist)4. Specific journals<ol style="list-style-type: none">a. Journal of Artificial Societies and Social Simulation (jasss.soc.surrey.ac.uk)b. Computational and Mathematical Organization Theory (www.springerlink.com/content/1381-298X)c. Complexity (www3.interscience.wiley.com/journal/388804/home) <p><u>E. Conference papers</u></p> <ol style="list-style-type: none">1. Winter Simulation Conference reports (www.wintersim.org) <p><u>F. Videos</u></p> <ol style="list-style-type: none">1. Google Video (www.google.com) <p><u>G. General reference</u></p> <ol style="list-style-type: none">1. Wikipedia (en.wikipedia.org)
---	---

F. SEARCH RESULTS

The following tables summarize the search results. From a review of more than 15,000 resources, I selected about 300 potentially useful resources (the “primary resources” and “referenced resources”). From these I selected about one hundred essential resources for inclusion in this report. Thus, of the total number of resources I reviewed, the essential resources represent about 1 percent.³

³ Because some of the essential resources came from sources other than the literature search (such as from recommendations of people who reviewed this report) the number of essential resources is greater than the sum of the ‘selected resources’ in the following table.

Complexity science – an Introduction for actuaries

Finding the hundred essential resources *continued*

F. SEARCH RESULTS CONTINUED

Search	Database	Search terms	Search focus	Results			
				Hits	Primary resources	Referenced resources	Selected resources
	<u>A. Actuarial online databases</u>						
A.1.a	SOA	CS	Any resources	41	9	11	5
A.1.b		ABMS	Any resources	48	1	0	1
A.2.a	CAS	CS	Any resources	12	3	1	1
A.2.b		ABMS	Any resources	29	3	13	1
A.3.a	IAA	CS	Any resources	31	3	3	3
A.3.b		ABMS	Any resources	77	2	0	0
A.4.a	Institute of Actuaries – UK	CS	Any resources	16	0	0	0
A.4.b		ABMS	Any resources	4	3	5	2
A.5.a	Institute of Actuaries – Australia	CS	Any resources	0	0	0	0
A.5.b		ABMS	Any resources	1	0	0	0
	<u>B. Books</u>						
B.1.a	Library of Congress	CS + Actuarial	Any text	331	12	42	12
B.1.b		ABMS + Actuarial	Any text	313	7	4	4
B.2.a	Google Books	CS + Actuarial	Any text	1,478	6	0	4
B.2.b		ABMS + Actuarial	Any text	5,132	3	0	1
B.3	Santa Fe Institute Book List	Manual search	All books	NA	1	0	0
	<u>C. Internet</u>						
C.1.a.i	Google Web	CS + Actuarial	PDF documents (top 200)	698,576	31	3	3
C.1.a.ii			non-PDF resources (top 200)	1,742,226	9	0	1
C.1.b.i		ABMS + Actuarial	PDF documents (top 200)	276,350	18	3	8
C.1.b.ii			non-PDF resources (top 200)	719,180	5	0	0

Complexity science – an introduction for actuaries

Finding the hundred essential resources *continued*

F. SEARCH RESULTS CONTINUED

Search	Database	Search terms	Search focus	Results			
				Hits	Primary resources	Referenced resources	Selected resources
<u>D. Journals</u>							
D.1.a	Google Scholar	CS		2,763	4	0	0
D.1.b		ABMS		8,100	3	0	0
D.2.a	PubMed	CS		58	9	0	1
D.2.b		ABMS		413	17	0	1
D.3	Santa Fe Institute Working Papers	Manual search	All working papers	NA	5	1	1
D.4.a	Journal of Artificial Societies and Social Simulation	Manual search	All articles	NA	15	0	3
D.4.b	Computational and Mathematical Organization Theory	Manual search	All articles 2008-2010	NA	2	0	0
D.4.c	Complexity	Manual search	All articles 2008-2010	NA	1	0	0
<u>E. Conferences</u>							
E.1	Winter Simulation Conference reports	Manual search	Relevant conference tracks 2008-2009	NA	6	0	3
<u>F. Videos</u>							
F.1	Google Video	CS	All videos	106	7	0	2
		ABMS	All videos	518	3	0	1
		Other	NA	NA	14	0	8
<u>G. General reference</u>							
G.1	Wikipedia	CS	All articles	2	2	0	0
		ABMS	All articles	1	1	0	0

GLOSSARY

actor: See ‘vertex’.

agent: The fundamental element of a Complexity Science model, representing an actor within a system.

agent-based model: A type of computer simulation that models the relationships and behaviors of agents within a complex system, in order to model the emergent behavior of the system as a whole.

artificial intelligence: A scientific field whose goal is to develop computers that think like humans.

artificial life: A scientific field that shows how computer programs can emulate certain features of living organisms.

asynchronous updating: When a behavior rule is not applied to all agents at one time step.

Barabási-Albert model: A type of network model.

behavioral economics: A relatively new field that experimentally investigates the behavior of humans in an economic system.

behavior rule: An algorithm governing how an agent’s states are updated from one time step to the next.

bond: See ‘edge’.

boundary: In reference to a cellular automaton in the form of a lattice graph: how cells at the edge of the lattice are related to other cells of the lattice.

catastrophe theory: A branch of mathematics popular in the 1970s that studies how large discrete changes (catastrophes) can appear in solutions of continuous equations with only small parameter changes.

cell: See ‘vertex’.

cellular automaton: A graph with vertices (cells) that can assume two or more states, and a behavior rule governing how each vertex’s state is updated.

characteristic path length: See ‘mean geodesic’.

chaos theory: The mathematical study of chaotic systems. See ‘chaotic system’.

chaotic system: A dynamic system that is highly sensitive to initial conditions. See ‘dynamic system’.

cluster: A group of vertices disconnected from other vertices in a graph, like an island.

complete N-graph: An undirected graph with a maximal number of edges.

complex adaptive system: A complex system that changes its behavior to respond to changes in its environment.

complex adaptive systems theory: One of the branches of Complexity Science, dealing with the theory and application of complex adaptive systems.

Complexity science – an introduction (and invitation) for actuaries

Glossary continued

complexity science: A new field that studies universal principles common to all complex systems.

complex system: An interesting dynamic system. See ‘dynamic system’.

complex systems actuary: A professional who addresses problems in complex systems, using the tools and concepts of Complexity Science.

computation: The aggregate behavior of a complex system, in which each of its agents carries out (or computes) its behavior rule, the way a computer carries out its program. Equivalently, the complex system can be said to be processing information. See ‘information’.

computational complexity theory: A branch of computer science that classifies computational tasks according to their inherent difficulty.

computational model: Agent-based model. See ‘agent-based model’.

connected triple: Three connected vertices (which may also be a ‘triangle’).

controlled experiment: An experiment that isolates the effect of one variable on a system by holding constant all variables but the one under observation.

correlation coefficient: A measure of the extent to which vertices are connected to other vertices with like degree.

criticality: A state where a system is out of balance, but not yet chaotic.

cybernetics: Originating in electrical engineering, this field studies the aggregate non-linear behavior of systems characterized by feedback loops.

degree: The number of edges connected to a vertex. A vertex of a directed graph has both an in-degree and an out-degree for each vertex, which are the numbers of in-coming and out-going edges respectively.

degree distribution: The distribution of vertices with various degrees in a graph. Usually depicted on a graph with the x axis representing the degrees, and the y axis representing the degree frequencies.

density: In a graph, the ratio of the number of edges to the possible number of edges.

design space: For a complex system, all possible combinations of relationships networks and behavior rules.

diameter: The length (in number of edges) of the longest of the geodesic paths between all pairs of vertices in a graph.

digraph: A directed graph.

directed edge: A edge that runs in only one direction between two vertices.

directed graph: A graph of which all the edges are directed.

dynamic system: A system together with a behavior rule that causes the state of at least one of its objects to change over time.

Complexity science – an introduction (and invitation) for actuaries

Glossary continued

dynamical systems theory: A branch of applied mathematics, this field studies systems that can be modeled with a particular class of mathematical equations (differential equations or difference equations).

dynamical network: A type of network that incorporates agent behavior rules .

edge: A line connecting two vertices. Also called a tie (sociology), a link (computer science), and a bond (physics).

edge of chaos: A description of the most productive complex adaptive systems, meaning that such systems are closer to random systems than to simple systems. Many question the concept's validity.

emergence: A characteristic of a complex systems, in which aggregate patterns arise out of the endogenous interactions of its agents with each other and an environment, without any central controller or other outside influence.

environment: An element of a complex system on which agents move and with which they can interact.

evolution: The creation of complex system behavior patterns that solve hard problems, most commonly problems of survival.

experimental mathematics: A branch of mathematics, this field uses computers and numerical computation to investigate mathematical objects and properties. Generally, the field studies objects and properties that have already been investigated using traditional mathematics.

first-order cellular automaton: A cellular automaton with a behavior rule that only depends on states at the previous time step. See 'second-order cellular automaton'.

fitness landscape: A 3-dimensional representation of a design space. See 'design space'.

fractal: A geometric pattern that is repeated at ever smaller scales to produce irregular shapes and surfaces that cannot be represented by classical geometry.

fractal geometry: A branch of mathematics that studies shapes in nature, and shows that many are not regular or smooth, but rather are nested shapes with intricate patterns.

fragile: Description of a system or network that cannot withstand damage to its parts.

game theory: A branch of applied mathematics that studies behavior in strategic situations where one person's (or organization's) choices depend on the choices of others.

general systems theory: A scientific field popular in the 1960s that studied general principles of social system functioning.

geodesic path: The shortest path from one vertex to another. There may be more than one geodesic path between two vertices. Also called, simply, the 'geodesic'. See also 'path length'.

graph: A representation of a real-world network, consisting of vertices and edges.

heterogeneous: Referring to agents of a complex system, when they can have different values for their attributes and states, different types of states, or different behavior rules.

heuristic: An informal guide to the solution of a problem. A rule of thumb.

Complexity science – an introduction (and invitation) for actuaries

Glossary continued

hub: In a graph, a vertex with a relatively high degree.

in-degree: The number of edges directed inward (with directional arrows) to a vertex.

information: The state of a complex system's environment, together with the states of all the agents, that are used to define the system's behavior rules.

lattice model: A type of network model.

layout: A pattern for placing vertices on a plot of a graph, determined by a placement algorithm.

link: See 'edge'.

loop: An edge that starts and ends on the same vertex.

mean degree: For a graph, the average of the degrees for all vertices.

mean geodesic: The average of all the geodesics of a graph. Also called the 'characteristic path length'.

minimum degree: In a graph, the degree of the vertex with the least degree.

Moore neighborhood: In a two-dimensional cellular automata, the eight cells surrounding a central cell. It is named after Edward F. Moore, a pioneer of cellular automata theory. It is one of the two most commonly used neighborhood types; the other is the four-cell von Neumann neighborhood. See 'von Neumann neighborhood'.

multi-agent model: Sometimes used synonymously with agent-based model. But the term can also mean a model in the field of multi-agent systems, a field mainly concerned with robot interactions. The term can also be used to mean a subset of agent-based models in which agents are heterogeneous.

multi-dimensional histogram: A histograms showing frequencies of more than one system state over time.

neighborhood: The collection of vertices connected to a vertex within a certain edge distance.

network: Any system that can be modeled using vertices to represent system elements or agents and using edges to represent relations or interactions among the elements.

network science: One of the branches of Complexity Science. It deals with the theory and application of networks. Also called network theory.

network theory: See 'network science'.

neural network: A virtual device, modeled after the human brain, in which several interconnected elements process information simultaneously, adapting and learning from past patterns.

node: See 'vertex'.

non-linear: A relationship that is not linear; that is, change in an independent variable may produce wildly non-proportional change in a dependent variable.

non-linear dynamics: A branch of mathematics, this field analyzes non-linear mathematical equations.

Complexity science – an introduction (and invitation) for actuaries

Glossary continued

object-oriented programming: A type of computer programming in which an ‘object’ includes both the object’s attributes (called ‘instance variables’) and the functions (called ‘methods’) that operate on the attributes.

oscillation: A common emergent pattern in a complex system, where its properties or attributes exhibit large swings.

out-degree: The number of edges directed outward (with directional arrows) from a vertex.

participatory model: An agent-based model in which people take agent roles and, following either simple scripted rules or their own instincts, act out the evolution of a simulation in the real world.

path length: The least number of edges between two vertices. Also called the ‘geodesic path length’.

pattern-matching algorithm: An algorithm that finds data patterns with potential near-term predictive value.

periodic boundary: A boundary where corresponding agents at opposite sides of a lattice are related as nearest neighbors.

phase-space diagram: Multi-dimensional diagrams that show the possible states of a system, with each state corresponding to a point on the diagram.

phase transition: A sudden dramatic change in a material from one state to another, such as ice turning to water, or non-magnetic material turning magnetic.

power law distribution: A degree distribution of many natural and human-made networks, conforming to the formula $p(X = x) \sim x^{-k}$. For phenomena following a power law, small occurrences are common, whereas large occurrences are rare.

punctuated equilibrium: A common emergent pattern of a complex system, in which the system goes through long periods of relative stasis, interspersed with brief periods of explosive activity.

random model: A type of network model.

random system: A dynamic system for which the state changes of its objects appear to be random. See ‘dynamic system’.

resilience: The capacity of a system or network to withstand damage to its parts.

robust: Description of a system or network that can withstand damage to its parts.

rulestring: A simple way to represent a behavior rule for a cellular automaton whose agents can have only two states, consisting of a string of 1’s and 0’s corresponding to the last column of the cellular automaton’s transition table

scale free network: A network that has a degree distribution that follows a power law. See ‘power law’.

second-order cellular automaton: A cellular automaton with a behavior rule that depends on states in the two previous time steps.

Complexity science – an introduction (and invitation) for actuaries

Glossary continued

self-organization: The propensity of dynamic systems to organize themselves into complex systems, on their own, without experimentation, mutation, or selection.

self-organized criticality: The tendency for a system to organize itself to a critical state, without outside manipulation.

serious game: A game whose primary purpose is training, education, or discovery. Serious games are also called ‘e-learning simulations’ and ‘simulation challenges’.

set: A collection of objects.

simple one-dimensional cellular automaton: A graph whose non-boundary vertices all have exactly two nearest neighbors (ie, neighbors within a radius of one).

simple system: A dynamic system for which the state changes of its objects are relatively uninteresting. See ‘dynamic system’.

simple two-dimensional cellular automaton: A graph whose non-boundary vertices all have either exactly four or exactly eight nearest neighbors.

site: See ‘vertex’.

skewed distribution: A non-symmetrical distribution with a long and often ‘fat’ tail. The power-law distribution is a skewed distribution.

small world effect: A characteristic of a network, whereby its mean geodesic is approximately equal to the mean geodesic of a random network with the same number of vertices and mean degree. Many real-world networks exhibit this effect.

small world network: A network for which (a) its mean geodesic is approximately equal to the mean geodesic of a similar random network (one with the same number of vertices and the same mean degree or, equivalently, number of edges), and (b) its transitivity is much greater than the transitivity of a similar random network. Many real-world networks are small world networks.

synchronous updating: When a behavior rule is applied to each agent at each time step.

system: A set whose objects are related to one another.

systems dynamics: Originating in electrical engineering, this field studies the aggregate non-linear behavior of systems characterized by feedback loops.

tie: See ‘edge’.

transition table: A method to present a set of simple if-then rules or a discrete function, whereby the values of the independent variables are listed in columns of a table, and the corresponding value of the dependent variable is listed in another column of the table.

transitivity: a measure of the probability that the adjacent vertices of a vertex are connected. It is equal to: $3 \times$ number of triangles in a graph/number of connected triples. See ‘triangle’ and ‘connected triple’. Also called the ‘clustering coefficient’, even though this term is also used for other graph measures.

triangle: Three vertices all of which are connected to each other.

Complexity science – an introduction (and invitation) for actuaries

Glossary continued

Turing Machine: A virtual machine consisting of a tape, a read/write head, and a transition table, that can perform any arithmetic or logical function.

undirected edge: An edge that runs in both directions between two vertices.

undirected graph: A graph of which all the edges are undirected.

universal computer: A Turing Machine that can simulate the behavior of any other Turing Machine, including itself. Also known as a Universal Turing Machine.

Universal Turing Machine: See ‘universal computer’.

vertex: The fundamental unit of a network. Also called a node (computer science), an actor (sociology), and a site (physics).

von Neumann neighborhood: In a two-dimensional cellular automata, the four cells orthogonally surrounding a central cell. It is named after John von Neumann, and is one of the two most commonly used neighborhood types; the other is the eight-cell Moore neighborhood. See ‘Moore neighborhood’.

Watts-Strogatz model: A type of network model.