Land Warfare and Complexity,
Part I: Mathematical
Background and Technical
Sourcebook (U)

Andrew Ilachinski

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

Center for Naval Analyses
4401 Ford Avenue • Alexandria, Virginia 22302-1498
Approved for distribution:

Dr. Igor Mikolic-Torreia
Director, Systems and Tactics Team
Operating Force Division

This document represents the best opinion of CNA at the time of issue.
It does not necessarily represent the opinion of the Department of the Navy.

CLEARED FOR PUBLIC RELEASE
Distribution limited to DOD agencies. Specific authority, WR37-003-2.
For copies of this document call: CNA Document Control and Distribution Section at 703-824-2943.

Copyright © 1996 The CNA Corporation
The purpose of this paper is to provide the theoretical framework and mathematical background necessary to understand and discuss the various ideas of nonlinear dynamics and complex system theory to plant seeds for a later, more detailed discussion (provided in Part II of this report) of how these ideas might apply to land warfare issues. This paper is also intended to be a general technical sourcebook of information. The main idea put forth in this paper is that significant new insights into the fundamental processes of land warfare can be obtained by viewing land warfare as a complex adaptive system. That is to say, by viewing a military "conflict" as a nonlinear dynamical system composed of many interacting semi-autonomous and hierarchically organized agents continuously adapting to a changing environment.
"So a military force has no constant formation, water has no constant shape: the ability to gain victory by changing and adapting according to the opponent is called genius."

– The Art of War, Sun Tzu (4th century B.C.)

"Strategy is a system of expedients. It is more than a science: it is the application of knowledge to practical life, the development of thought capable of modifying the original guiding idea in the light of ever-changing situations; it is the art of acting under the pressure of the most difficult conditions."

– Helmuth von Moltke (1800-1891)

"Everything is very simple in war, but the simplest thing is difficult. These difficulties accumulate and produce a friction, which no man can imagine exactly who has not seen war...This enormous friction, which is not concentrated, as in mechanics, at a few points, is therefore everywhere brought into contact with chance, and thus facts take place upon which it was impossible to calculate, their chief origin being chance."

– Carl von Clausewitz (1780-1831)

"Like friction and uncertainty, fluidity is an integral attribute of the nature of war. Each episode in war is the temporary result of a unique combination of circumstances, requiring an original solution. But no episode can be viewed in isolation. Rather, each merges with those that precede and follow it—shaped by the former and shaping the conditions of the latter—creating a continuous, fluctuating fabric of activity replete with fleeting opportunities and unforeseen events. Success depends in large part on the ability to adapt to a constantly changing situation."

"The occurrences of war will not unfold like clockwork. Thus, we cannot hope to impose precise, positive control over events. The best we can hope for is to impose a general framework of order on the disorder, to prescribe the general flow of action rather than to try to control each event."

– Warfighting, FMFM-1
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overview</strong></td>
<td>1</td>
</tr>
<tr>
<td>Two Intriguing Questions</td>
<td>1</td>
</tr>
<tr>
<td>Two Words of Caution</td>
<td>3</td>
</tr>
<tr>
<td>Infancy</td>
<td>3</td>
</tr>
<tr>
<td>Buzzwords</td>
<td>3</td>
</tr>
<tr>
<td>Organization of Paper</td>
<td>5</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td>7</td>
</tr>
<tr>
<td>Is it time for a fresh new look at land warfare modeling?</td>
<td>7</td>
</tr>
<tr>
<td>A Proposal</td>
<td>8</td>
</tr>
<tr>
<td>Complex Systems</td>
<td>9</td>
</tr>
<tr>
<td>What Complex Systems Theory Is</td>
<td>14</td>
</tr>
<tr>
<td>What Complex Systems Theory Is Not</td>
<td>16</td>
</tr>
<tr>
<td>Land Combat as a Complex Adaptive System?</td>
<td>16</td>
</tr>
<tr>
<td>Redefined Conventions</td>
<td>19</td>
</tr>
<tr>
<td><strong>Nonlinear dynamics and Chaos</strong></td>
<td>21</td>
</tr>
<tr>
<td>Introduction</td>
<td>21</td>
</tr>
<tr>
<td>Short History</td>
<td>21</td>
</tr>
<tr>
<td>Dynamical Systems</td>
<td>23</td>
</tr>
<tr>
<td>Discrete-time Poincare maps</td>
<td>24</td>
</tr>
<tr>
<td>Phase Space Volumes</td>
<td>25</td>
</tr>
<tr>
<td>Dissipative Dynamical Systems</td>
<td>25</td>
</tr>
<tr>
<td>Strange Attractors</td>
<td>26</td>
</tr>
<tr>
<td>Deterministic Chaos</td>
<td>26</td>
</tr>
<tr>
<td>Conservative Dynamical Systems</td>
<td>27</td>
</tr>
<tr>
<td>Example #1: The Bernoulli Shift Map</td>
<td>28</td>
</tr>
<tr>
<td>Stability</td>
<td>29</td>
</tr>
<tr>
<td>Predictability</td>
<td>29</td>
</tr>
<tr>
<td>Deterministic Randomness</td>
<td>30</td>
</tr>
<tr>
<td>Computability</td>
<td>30</td>
</tr>
<tr>
<td>Example #2: The Logistic Map</td>
<td>31</td>
</tr>
<tr>
<td>Definition</td>
<td>32</td>
</tr>
<tr>
<td>Fixed Point Solutions</td>
<td>33</td>
</tr>
<tr>
<td>Universality</td>
<td>34</td>
</tr>
<tr>
<td>Behavior for</td>
<td>35</td>
</tr>
<tr>
<td>Two-Dimensional Strange Attractors</td>
<td>37</td>
</tr>
<tr>
<td>Henon Map</td>
<td>37</td>
</tr>
<tr>
<td>Qualitative Characterization of Chaos</td>
<td>39</td>
</tr>
<tr>
<td>Time Dependence</td>
<td>39</td>
</tr>
<tr>
<td>Poincare Maps</td>
<td>40</td>
</tr>
<tr>
<td>Autocorrelation Function</td>
<td>40</td>
</tr>
<tr>
<td>Power Spectrum</td>
<td>41</td>
</tr>
<tr>
<td>Quantitative Characterization of Chaos</td>
<td>42</td>
</tr>
</tbody>
</table>
Lyapunov Exponents ........................................ 42
Information Theoretic Interpretation .................. 44
Numerical Computation ................................. 44
Fractal Dimensions ........................................ 46
Box Dimension ........................................... 46
Information Dimension ................................. 48
Correlation Dimension .................................. 49
Hierarchy of Generalized Fractal Dimensions ...... 50
Lyapunov Dimension ...................................... 50
Kolmogorov-Sinai Entropy ............................... 51
Time-Series Forecasting and Predictability .......... 52
State-Space Reconstruction via Embedding .......... 53
Chaotic Control ......................................... 54
Brief Overview of Method .............................. 56
Lessons of Nonlinear Dynamics and Chaos .......... 57
Tools for the Decision Maker? ........................... 59

Complex systems .......................................... 61
Introduction ............................................. 61
Short History ........................................... 62
Ants and brains ... and combat forces? ................. 63
Collectivism ............................................. 64
Self-Organization ....................................... 66
Emergence .............................................. 67
Example #1: Reynolds's Boids ......................... 67
Example #2: Collective Decentralized Sorting ..... 68
Edge-of-Chaos .......................................... 70
Complexity as a Measure? ............................. 72
What is the Difference Between Chaos and Complexity? 74
Cellular Automata ....................................... 74
Example #1: One-dimensional CA ....................... 75
Example #2: Conway's Life ............................ 80
Example #3: Belousov-Zhabotinski Reaction ........ 82
Example #4: Lattice Gases ............................. 83
Example #5: Collective Behavior in Higher Dimensions ................................................. 84
Other Variants ......................................... 85
Genetic Algorithms ..................................... 85
Genetic Operators ...................................... 86
The Basic GA Recipe .................................. 87
Example #1: Function Maximization ................ 88
Example #2: Local Forecasting of High-Dimensional Chaotic Dynamics ..................... 90
Other Variants ......................................... 91
Self-Organized Criticality ............................. 92
Example: Sandpiles .................................... 94
Fractals .................................................. 96
1/f-Noise .................................................. 96
A Possible Connection with Land Combat? ................. 96
Complex Adaptive Systems .................................. 97
Characteristics ................................................ 99
Agent-Based Simulations .................................... 100
Adaptive Autonomous Agents ............................... 101
What Distinguishes the Study of Adaptive
Autonomous Agents from Traditional AI? .............. 102
What Distinguishes Traditional Modeling
Approaches from Agent-Based Simulations? ........... 103
Swarm ......................................................... 104
Neural Networks ............................................. 105
A Short History ............................................. 107
An Heuristic Discussion .................................. 108
Defining and Training a Neural Network ................. 109
Using a Trained Net ........................................ 111
General Model Development: A Short Primer .......... 112
Backpropagation Algorithm ................................ 113
Pseudo-Code ................................................ 114
Example: NETalk .......................................... 115
Other Designs .............................................. 117
Lessons of Complex Systems Theory ...................... 118

Land Warfare and Complexity Theory: Preliminary

Musings ........................................................ 121
Framing the Problem ....................................... 123
Chaos in combat models .................................... 125
Non-Monotonicities and Chaos ............................ 126
Minimalist Modeling ....................................... 127
Generalizations of Lanchester's equations .............. 128
Demonstration of chaos in war using historical data .... 129
Nonlinear dynamics and chaos in arms-race models .... 130
Combat simulation using cellular automata .............. 130
Computer viruses ("computer counter-measures") ........ 131
Intelligent Software Agents ................................ 132
Agent-based simulations ................................... 133
Tactics and strategy evolution using genetic
algorithms ...................................................... 136
Time-series analysis ........................................ 137
Relativistic information .................................... 137
Exploitation of Characteristic Time Scales .............. 138

Summary and Conclusion .................................... 141
Basic Concepts ............................................. 141
Mathematical Tools ........................................ 141
Basic Lessons Learned .................................... 143
Nonlinear Dynamics ........................................ 143
Complex Systems .......................................... 144
Appendix A: World Wide Web Nonlinear Dynamics and Complex Systems Theory Resources

Subject-Sorted WWW Link Listing

General Sources .................................................. 150
Artificial Intelligence ........................................... 150
Artificial Life ...................................................... 151
Artificial Life Simulation and Research Groups .............. 151
Autonomous Agents ............................................... 152
Cellular Automata .................................................. 152
Chaos ................................................................. 153
Fractals ............................................................... 153
Fuzzy Logic ......................................................... 154
Genetic Algorithms ................................................. 154
Genetic Programming .............................................. 155
Intelligent Software Agents ...................................... 155
Neural Nets ........................................................ 155
Nonlinear Dynamics ................................................ 156
Software ............................................................. 156
Time Series Analysis ................................................ 157
Alphabetized WWW Link Listing in HTML format .......... 158

Appendix B: Glossary of Terms

Adaptation .......................................................... 179
Algorithmic Complexity .......................................... 179
Animats ............................................................... 179
Artificial Life ....................................................... 179
Attractor .............................................................. 179
Autonomous (or Adaptive-) Agent .............................. 179
Autopoietic Systems .............................................. 180
Autopoiesis .......................................................... 180
Backpropagation Algorithm ..................................... 181
Basin of Attraction ............................................... 181
Bifurcation .......................................................... 181
Boolean Function .................................................. 181
Cantor Set ........................................................... 181
Catastrophe Theory ................................................ 182
Cellular Automata .................................................. 182
Cellular Games ..................................................... 182
Chaos ................................................................. 182
Chaotic Control ..................................................... 183
Classifier Systems ................................................ 183
Class-P Problems .................................................. 183
Co-Adaptation/Co-Evolution ................................... 183
Complex Adaptive Systems ..................................... 184
Complexity ........................................ 184
Computational Complexity .................. 184
Computational Irreducibility .......... 185
Computational Universality ........... 185
Conservative Dynamical Systems .... 185
Cost Function ................................. 185
Coupled-Map Lattices ....................... 186
Criticality .................................... 186
Crossover Operator ......................... 186
Dissipative Structure ...................... 186
Edge-of-Chaos ................................ 186
Emergence .................................. 187
Entropy ...................................... 187
Ergodic System ................................ 187
Ergodic Theory ................................ 187
Evolution ................................... 188
Evolutionary Programming ............ 188
Evolutionary Stable Strategy ....... 188
Finite Automata .............................. 188
Fitness Landscape ......................... 188
Flicker- (or 1/f) Noise .................... 188
Fractals ..................................... 189
Fractal Dimension ......................... 189
Frustration .................................. 189
Fuzzy Logic .................................. 189
Genetic Algorithms ......................... 189
Genotype .................................... 190
Hamiltonian System ....................... 190
Hausdorff Dimension ...................... 190
Hierarchy .................................... 190
Homoclinic Point ............................ 190
Hopf-Bifurcation ...................... 190
Hypercyle ................................... 190
Information Dimension ................. 190
Information Theory ...................... 191
Intermittency ................................ 191
Kolmogorov Entropy ....................... 191
Lattice Gas Models ......................... 191
Life Game .................................... 192
Limit-Cycle .................................. 192
Lindenmeyer (or L-) Systems ........ 192
Logistic Equation ......................... 192
Lotka-Volterra Equations ............... 193
Lyapunov Exponent ....................... 193
Markov Process .............................. 193
Maximum Entropy ........................... 194

V
Mean-Field Theory ........................................ 194
Multifractal ............................................... 194
Navier-Stokes Equations ................................. 194
Neural Networks .......................................... 194
Nonlinearity ................................................ 195
NP-Hard Problems ........................................ 195
NP-Complete ............................................. 195
Order Parameter .......................................... 195
Percolation Theory ....................................... 195
Petri Nets .................................................. 196
Phase Space ............................................... 196
Phase Transition ......................................... 196
Phenotype ................................................. 196
Poincare Map ............................................. 196
Prisoner's Dilemma ....................................... 197
Probabilistic CA .......................................... 197
Punctuated Equilibrium ................................. 197
Quasiperiodic ............................................. 197
Random Boolean Networks ............................. 197
Reaction-Diffusion Models ............................. 198
Relativistic Information Theory ....................... 198
Scaling Laws ............................................. 198
Search Space ............................................. 199
Self-Organization ....................................... 199
Self-Organized Criticality ............................ 199
Simulated Annealing .................................... 199
Solitons .................................................... 199
Spatio-Temporal Chaos ................................. 199
Spin Glasses .............................................. 200
Strange Attractors ...................................... 200
Symbolic Dynamics ...................................... 200
Synergetics .............................................. 200
Topological Dimension ................................ 200
Universality .............................................. 201
Unstable Equilibrium .................................. 201

Appendix C: Recommended Reading ................... 203
References ............................................... 205
Overview

The purpose of this paper is to provide the theoretical framework and mathematical background necessary to understand and discuss the various ideas of nonlinear dynamics and complex systems theory and to plant seeds for a later, more detailed discussion (that will be provided in Part II of this report\(^1\)) of how these ideas might apply to land warfare issues. This paper is also intended to be a general technical sourcebook of information.

Two Intriguing Questions

Question 1: What does the behavior of the human brain have in common with what happens on a battlefield?

The human brain is composed of about ten billion neurons, each of which, on average, is connected to about a thousand other neurons. What each neuron does is a complicated function of what it did before and what its thousand or so neighbors were doing. Somehow, mysteriously, for reasons that are still not quite clear and perhaps never will be fully, this cauldron of ceaseless neuro-chemical activity spawns something called "consciousness" that emerges on a much higher level than the one on which any of the brain's constituent parts themselves live.

Nowhere is there a prescription for an "awareness of self." Nowhere is there a hard-wired rule that says this arrangement of neurons will prefer football to boxing and that arrangement will prefer soccer to both. Nowhere on the neuronal level is there a rule that assigns the personality that is uniquely mine. These are all emergent, higher-level phenomena that, while owing their existence to the myriad interactions of ten billion neurons, cannot be deduced directly from them. As such, the human brain is the prototypical example of a complex system, or a system composed of many nonlinearly interacting parts. Now, what happens on a battlefield?

While no battlefield can possibly consist of as many combatants as there are neurons in a human brain, the analogy between what makes the human brain "interesting" and what makes that which happens on a battlefield "complicated" is not such a poor one. Both consist of a large number of nonlinearly interacting parts whose individual behavior depends on the action and pattern of behavior of other (nearby and not-so-nearby) parts, both obey a decentralized control, both appear to be locally "chaotic" but harbor long-range order, both tend not to dwell for long times near equilibrium, preferring instead to

---

\(^1\) *Land Warfare and Complexity, Part II: An Assessment of the Applicability of Nonlinear Dynamics and Complex Systems Theory to the Representation of Land Warfare* is scheduled to be delivered to sponsor for review 1 July, 1996.
exist almost exclusively in a nonequilibrium state, and both must continually adapt to internal and external pressures and to the environment. On paper, at least, the human brain and the battlefield appear to have much in common.

Question 2: Might there be higher-level processes that emerge on the battlefield, in the way consciousness emerges in a human brain?

Both of these enormously difficult, yet intriguing, questions are clearly meant to be taken rhetorically (at least for the moment). However, the motivation behind asking these two questions is what lies at the heart of this report and the overall project of which it is a part. The goal of this paper is to provide the reader with a basic set of tools and concepts out of which a tentative real answer to this question might conceivably (some day) emerge. A hint that these questions are neither ill-posed nor simply foolish is provided by an emerging new field of research that can loosely be called complex systems theory.

In recent years there has been a rapid growth in what has come to be popularly known as the New Sciences of Complexity. Despite being somewhat of a misnomer, because the "science" is arguably more a philosophy of looking at behaviors of complex systems than a rigorous well-defined methodology, this emerging field nonetheless has many potentially important new insights to offer into the understanding of the behaviors of complex systems.

A complex system can be thought of, generically, as any dynamical system composed of many simple, and typically nonlinearly, interacting parts. A complex adaptive system is a complex system whose parts can evolve and adapt to a changing environment. Complex systems theory is then the study of the behavior of such systems, and is rooted in the fundamental belief that much of the overall behavior of diverse complex systems — such as natural ecologies, fluid flow, the human brain, the Internet, perhaps even on the battlefield, etc — in fact stems from the same basic set of underlying principles.

While research in this still-developing field has yet to produce an all-encompassing "theory of complexity," it has already introduced promising new analytical methodologies and has uncovered many provocative and useful organizing principles. The fundamental challenge of this study is to assess what insights into the understanding of land warfare can be gained by using the tools and methodologies developed for the general study of nonlinear dynamics and complex systems.
Two Words of Caution

Before we begin the discussion in earnest, two important words of caution are in order. The two words are *infancy* and *buzzwords*.

*Infancy*

Remember that, at the present time, complexity cannot be regarded as anything but an *infant science*! The *Santa Fe Institute* in New Mexico, for example, which is widely recognized as being a leading research center for complex systems, was founded just a decade ago in 1984. Moreover, many of the analytical tools and models developed for the study of complex systems, such as genetic algorithms and agent-based simulations, have been either developed or refined as part of the artificial-life research effort that itself sprang up only in 1987. Consequently, it would be premature – and unfair to complex systems theory – to expect to find a mature set of tools and methodologies at such an early stage of this burgeoning field's development.

Indeed, it is fair to say that it is as difficult to predict the potentially deep and lasting implications of research in complex systems theory as it is understood and practiced *today* as it would have been difficult to predict the implications of the state-of-the-art in, say, thermodynamics as it was understood circa 1820. As was true of thermodynamics then, 60 or so years prior to its full maturation, it is true to say of complex systems theory *now*, that the "killer application" (as it is often called in commercial software circles) or the "killer insight" (as it is sometimes called in physics) has not yet been born. We stress that all speculations about possible applications of complex systems theory, whether they appear in this report or elsewhere, must be interpreted in this light.

*Buzzwords*

Many buzzwords have appeared in the popular literature in recent years, not all of which have been described accurately. Terms and concepts such as *new sciences, chaos, complexity, complex systems theory, complex adaptive systems*, and so on, are commonly used to denote the ostensibly same fundamental core of principles.

In fact, there is no universally agreed upon definition of complex systems theory. Complex systems theory is a catch-all phrase that embodies a remarkably wide variety of disciplines ranging from biology, chemistry, and physics to anthropology to sociology to economics. Its many subfields include (but are not limited to) nonlinear dynamics, artificial life, evolutionary and genetic programming, cellular automata cellular games, agent-based modeling, among many others.
Table 1. A small sampling of research areas, concepts and tools all falling under the broad rubric of "new sciences"

<table>
<thead>
<tr>
<th>Research Areas</th>
<th>Concepts</th>
<th>Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>agent-based simulations</td>
<td>adaptation</td>
<td>agent-based simulations</td>
</tr>
<tr>
<td>artificial life</td>
<td>autonomous agents</td>
<td>backpropagation</td>
</tr>
<tr>
<td>catastrophe theory</td>
<td>autopoiesis</td>
<td>cellular automata</td>
</tr>
<tr>
<td>cellular automata</td>
<td>complexity</td>
<td>cellular games</td>
</tr>
<tr>
<td>cellular games</td>
<td>computational irreducibility</td>
<td>chaotic control</td>
</tr>
<tr>
<td>chaos</td>
<td>computational universality</td>
<td>entropy</td>
</tr>
<tr>
<td>chaotic control theory</td>
<td>criticality</td>
<td>evolutionary programming</td>
</tr>
<tr>
<td>complex adaptive systems</td>
<td>dissipative structures</td>
<td>fuzzy logic</td>
</tr>
<tr>
<td>coupled-map lattices</td>
<td>edge-of-chaos</td>
<td>genetic algorithms</td>
</tr>
<tr>
<td>discrete dynamical systems</td>
<td>emergence</td>
<td>inductive learning</td>
</tr>
<tr>
<td>evolutionary programming</td>
<td>fractals</td>
<td>information theory</td>
</tr>
<tr>
<td>genetic algorithms</td>
<td>intermittency</td>
<td>Kolmogorov entropy</td>
</tr>
<tr>
<td>lattice-gas models</td>
<td>phase space</td>
<td>lattice-gas models</td>
</tr>
<tr>
<td>neural networks</td>
<td>phase transitions</td>
<td>Lyapunov exponents</td>
</tr>
<tr>
<td>nonlinear dynamical systems</td>
<td>prisoner's dilemma</td>
<td>maximum entropy</td>
</tr>
<tr>
<td>percolation theory</td>
<td>punctuated equilibrium</td>
<td>neural networks</td>
</tr>
<tr>
<td>petri nets</td>
<td>self-organization</td>
<td>Poincare maps</td>
</tr>
<tr>
<td>relativistic information theory</td>
<td>self-organized criticality</td>
<td>power spectrum</td>
</tr>
<tr>
<td>self-organized criticality</td>
<td>strange attractors</td>
<td>symbolic dynamics</td>
</tr>
<tr>
<td>time-series analysis</td>
<td>synergetics</td>
<td>time-series analysis</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
</tbody>
</table>

Table 1 partitions many of the more common terms that one often hears in connection to "complex system theory" into three categories:

- **Research areas**, representing large (often interdisciplinary) fields of current research. This means, in particular, that a great deal remains to be learned about almost all of the areas listed in table 1. For example, while genetic algorithms are undoubtedly useful and powerful tools, there are also a large number of difficult open-ended problems that researchers are trying to solve about some of their most basic behaviors.

- **Concepts**, denoting the set of ideas, conjectures, hypotheses, organizing principles, and so on, that complex systems theory has so far spawned. It represents a sampling of the vocabulary of "The New Sciences" and appears in table 1 to remind the reader that these terms are often misunderstood, misused, or incorrectly defined altogether.

- **Tools**, representing the practical set of working methodologies, qualitative and quantitative measures, and mathematical descriptions that researchers have found to be useful in describing the properties and behaviors of complex systems.
Note that there is some overlap among the three categories appearing in table 1, particularly between the listings appearing under Research Areas and Tools. This underscores the fact that while nonlinear dynamics and complex systems theory have both amassed an impressive arsenal of practical and theoretical tools, these sets of tools are still evolving and are the subject of ongoing research. Notice also that there is an important, and ironic, feedback lurking here. The very tools that can, and should, be used to explore as large a variety of complex systems as possible — including, as suggested by this report, the processes that take place on a real battlefield — are themselves continuously refined and redefined in the process of studying those systems. Agent-based simulations, for example, virtually define much of what makes up artificial life studies — so that they are undeniably powerful tools with many important insights to offer — yet there is much that is still not understood about their basic design and how to make the best use of them. Similarly, while neural net technology has been around long enough that many powerful products have become available on the commercial market, one should not lose sight of the fact that researchers are still exploring the best way to make use of this new technology.

The point of this cautionary discussion is simply to dispel any prior notions that complex systems theory is nothing but a "canned set of software routines" sitting on a shelf somewhere, ready to be installed on a PC or MAC and used on whatever "complex problem" happens to strike one's fancy. The reality is that complex systems theory is, at the present time, a not terribly-well-defined and very much evolving set of ideas and tools with which the many complex systems appearing in nature can be studied. What makes complex systems theory an interesting prospect to mine for ideas about what may or not really be happening on the battlefield, is that in its short life it has already managed to produce an impressive list of useful insights into the understanding of the general behavior of complex systems. The goal of this project is to explore whether this list of insights also extends to the military realm, and particularly to land combat.

**Organization of Paper**

The paper begins with a general heuristic discussion of what is meant by a complex system and why it represents a fundamentally new approach to land warfare.

The discussion in the next section is more formal and is a mathematical introduction to nonlinear dynamics and deterministic chaos. This discussion summarizes the pertinent vocabulary, and uses simple examples to illustrate the basic ideas. Overviews of both qualitative and quantitative characterizations of chaos, as well as a discussion of relatively recent advances in chaotic control theory and
attractor-reconstruction from experimental time-series are also included. This section concludes with some important lessons learned from nonlinearity and chaos and explains how these lessons can be used by the decision maker.

The next section focuses on complex systems, beginning with a short history and simple examples of the central concept of emergence. Some of the more important tools of complex systems theory – including cellular automata, genetic algorithms, neural networks, and agent-based simulations – are discussed next. Attention is also given to an important idea called self-organized criticality, which is arguably the only existing holistic mathematical theory of self-organization in complex systems. The section concludes with some general lessons learned from complex systems theory.

The next section provides some preliminary musings on the possible applicability of nonlinear dynamics and complex systems theory to the understanding and/or representation of land warfare. The discussion in this section is intended to be a brief overview of the in-depth analysis and discussion of ideas that will be provided in Part II of this report.

The main points of this paper are summarized in the conclusion.

The two appendices and References provide additional reference material and may be consulted as information sourcebooks. Appendix A provides both a brief subject-sorted listing of information sources currently available on the World Wide Web – consisting of 91 URL links sorted into 16 categories – and an unsorted but much more extensive alphabetized listing of approximately 700 URL links. Appendix B provides a glossary of 100 terms commonly used in nonlinear dynamics and complex systems theory. An extensive reference list consisting of 330 journal articles, conference proceedings, monographs, texts and popularizations appears at the end of the paper.
Introduction

"...war is not an exercise of the will directed at inanimate matter, as in the case with the mechanical arts, or at matter which is animate but passive and yielding as in the case with the human mind and emotion in the fine arts. In war, the will is directed at an animate object that reacts."

– Carl von Clausewitz, On War

Is it time for a fresh new look at land warfare modeling?

In 1914, F. W. Lanchester [183] introduced a set of coupled ordinary differential equations as models of attrition in modern warfare. Similar ideas were proposed around that time by Chase [51] and Osipov [286].

While Lanchester's equations capture some important elements of combat, they are applicable only under a strict set of assumptions. These include having homogeneous forces that are continually engaged in combat, firing rates that are independent of opposing force levels and are constant in time, and units that are always aware of the position and condition of all opposing units, among many others. Lanchester's equations also suffer from a number of significant shortcomings, including modeling combat as a deterministic process, requiring knowledge of "attrition-rate coefficients" (the values of which are, in practice, very difficult if not impossible to obtain), an inability to account for any suppressive effects of weapons, an inability to account for terrain effects, and the inability to account for any spatial variation of forces. Generally speaking, Lanchester's equations simply lack the spatial degrees-of-freedom to model real-world combat. More importantly, they also leave out the all-important human factor – that is, the psychological and/or decision-making capability of the human operator.

While there have been many extensions to and generalizations of Lanchester's equations over the years (see discussion on page 139), very little has really changed in the way we fundamentally view and model combat attrition. It is a bit ironic that in this modern age of distributed interactive simulation and gigabyte-sized code driving networked 3D virtual-reality systems with embedded artificial intelligence, the underlying principles of combat attrition calculations in land warfare models are still largely the same as they were in Lanchester's time; this, despite the acknowledged deficiencies of Lanchester's equations. The question is, "Is there anything better?" Is there a better way – perhaps a way that bucks convention – of modeling land combat?

Recent developments in nonlinear dynamics and complex systems theory provide a potentially powerful new set of theoretical and practical tools to address many of the deficiencies mentioned above.
These developments also potentially represent a fundamentally new way of looking at land combat.

It is not an accident that the Lanchester equations have essentially the same mathematical form as the equations used for studying predator-prey relationships in natural ecologies of competing species. They both describe systems that evolve according to more or less the same basic driving forces of attrition. But while biologists have long known that there are universal patterns of behavior in the evolution of ecologies that transcend the closed-form equations that are typically used to model their behavior, the same thinking has not yet strongly influenced ground warfare modeling. We propose that there are also universal patterns of behavior underlying and driving the evolution of military combat that transcend the grossly simplified form of the Lanchester equations and the approximations and assumptions that they embody.

A Proposal

The main idea put forth in this paper is that significant new insights into the fundamental processes of land warfare can be obtained by viewing land warfare as a complex adaptive system (CAS). That is to say, by viewing a military "conflict" as a nonlinear dynamical system composed of many interacting semi-autonomous and hierarchically organized agents continuously adapting to a changing environment.

Compared to most conventional modeling philosophies, this approach represents a fundamental shift in focus:

from....

"Hard-wiring" into a model a sufficient number of (both low- and high-level) details of a system to yield a desired set of "realistic" behaviors – the rallying cry of such models being "more detail, more detail, we need more detail!"

...to....

Looking for universal patterns of high-level behavior that naturally and spontaneously emerge from an underlying set of low-level interactions and constraints – the rallying cry in this case being "allow evolving global patterns to emerge from the local rules!"

We should emphasize from the start what this study is and what it is not. This study is not a wholesale attempt to replace all Lanchester equation-based modeling of combat attrition. Just as the Lotka-Volterra equations for predator-prey dynamics capture some of the essence of population dynamics in natural ecologies, the Lanchester equations, or
one of their countless generalizations, may well describe combat attrition under an appropriate set of combat conditions. This study instead proceeds on the much broader charter of identifying the tools and methodologies that have been developed for the general study of nonlinear dynamics and complex systems and adapting them – along with whatever insights they might provide – to the modeling of land warfare.

To be more specific, two relatively new mathematical modeling techniques are applied to the modeling of land warfare as a complex adaptive system:

- local-rule-based dynamics patterned after cellular automata models
- parameters of the local decision space and the formulation of strategy and/or tactics patterned after genetic algorithms

Cellular automata and genetic algorithms are both common tools in the repertoire of tools used to describe and study complex systems. A self-contained discussion of what they are and how they are applied appears in the main text of this paper. The proposed methodology also makes fundamental use of game theory and neural networks.

Before discussing how land combat may be viewed as a complex adaptive system, it is prudent to first introduce the general notion of complex system.

**Complex Systems**

Consider some familiar examples of dynamics in complex systems:

- the predator-prey relationships of natural ecologies
- the economic dynamics of world markets
- the chaotic dynamics of global weather patterns
- the firing patterns of neurons in a human brain
- the information flow on the Internet
- the apparently goal-directed behavior of an ant colony
- the competing strategies of a nation's political infrastructure
many, many others.....

Now consider what these systems all have in common. Apart from obviously evolving according to very complex dynamics, almost all complex systems also share these fundamental properties:

- **Complex systems consist of—and their overall behavior stems from—a large assemblage of interconnected (and typically nonlinearly) interacting parts**

  - each of the systems listed above, as well as countless other examples of complex systems that one can write down, owe their apparent complexity to the fact that they consist not just of parts, but of parts whose states continually change as a function of the continual changes undergone by other parts to which they are connected.

- **Complex systems tend to be organized hierarchically, with complex behavior arising from the interaction among elements at different levels of the hierarchy**

  - whether we consider galactic systems, living organisms, or social or military organizations, the various structures making up all such systems are almost always organized in a hierarchy
    - individual parts of systems (usually called agents) form higher-level groups that act as agents that can then interact with other agents; these groups, in turn, form super-groups that also act as agents, interacting with other agents (though perhaps on a different timescale); and so on.
  
  - moreover, every part of the hierarchy is driven by two opposite tendencies:
    - an **integrative tendency**, compelling it to function as a part of the larger whole (on higher levels of the hierarchy)
    - a **self-assertive tendency**, compelling it to preserve its individual autonomy

- **The overall behavior of complex systems is self-organized under a decentralized control**

  - there is no God-like "oracle" dictating what each and every part ought to be doing, no master "neuron" telling each neuron when and how to "fire"
○ parts act locally on local information and global order emerges without any need for external control

○ According to Stuart Kauffman, who is one of the leading researchers of complex systems theory, "contrary to our deepest intuitions, massively disordered systems can spontaneously 'crystallize' a very high degree of order." [171]

○ self-organization takes place as a system reacts and adapts to its external environment

• **Overall behavior is emergent**

○ the properties of the "whole" are not possessed by, nor are they directly derivable from, any of the "parts"

○ examples: a line of computer code cannot calculate a spreadsheet, an air molecule is not a tornado and a neuron is not conscious

○ emergent behaviors are typically novel and completely unanticipated

○ elements of emergent behaviors may be universal, in the sense that more than one local rule set may induce more or the less same global behavior

• **Long-term behavior typically consists of a nonequilibrium order**

○ nonequilibrium order refers to organized states (sometimes called dissipative structures) that remain stable for long periods of time despite matter and energy continually flowing through them

○ a vivid example of nonequilibrium order is the Great Red Spot on Jupiter (see figure 1). This gigantic whirlpool of gases in Jupiter's upper atmosphere has persisted for a much longer time (on the order of centuries) than the average amount of time any one gas molecule has spent within it

• **Parts consist more of niches that need to be filled rather than of distinct labeled entities that carry an importance all their own**

○ the importance of a given "part" is dictated more by how that part interacts with the whole – and the "meaning" that particular interaction or set of interactions has in context of
the whole – than by what that part represents apart from the whole

○ as an example, consider how human beings now support themselves by driving cars, typing away at computer keyboards, sending faxes cross-country... activities that were nonexistent at the turn-of-the-century; moreover, it does not matter who repairs a malfunctioning computer, but only that the computer repair service exists

Figure 1. The Great Red Spot on Jupiter

○ Behavior cannot be described by reductionist methods alone

○ The traditional Western scientific method is predicated on a reductionist philosophy, in which the properties of a system are deduced by decomposing the system into progressively smaller and smaller pieces. However, in so doing, the emergent properties of a system are lost. In the act of exploring properties, reductionism loses sight of the dynamics. The analysis of complex systems instead requires a holistic, or constructionist, approach.

○ the properties of entities occupying "niches" on high-levels of the hierarchy influence, in a nonlinear fashion, entities occupying "niches" on the lower levels of the hierarchy; properties of these low-level entities in turn feed back up to influence the behavior of the high-level entities – the lifeblood of all complex adaptive systems is this continuous cycling of information from top to bottom to top to bottom...

○ the complex systems theory "approach" is sometimes concisely referred to as collectivist, a term designed to distinguish it from more traditional – and uni-directional –
topdown (or purely reductionist) or bottom-up (or purely synthetist) approaches

- Structure- and process-driven dynamics *vice simple aggregate of individual parts of a structure*

  - the essence of a complex system is its continual adaptive evolution; no static picture – such as is obtained, say, by simply listing a system's parts and how they interact – can adequately capture the often latent and otherwise subtle patterns that such systems tend to exhibit over long times

  - complex systems theory research consists not so much of observing what state a system happens to be in at what time as observing what kinds of patterns of behavior systems exhibit over the course of their entire evolution

One of the basic questions begging to be asked of all complex systems is, *"What are the universal patterns of behavior?"* According to thermodynamics and statistical mechanics, the critical exponents describing the divergence of certain physical measurables (ex: specific heat, magnetization, correlation length, etc.) are *universal* at a phase transition in that they are essentially independent of the physical substance undergoing the phase transition and depend only on a few fundamental parameters (such as the dimension of the space and the symmetry of the underlying order parameter). In like manner, an important driver fueling the fervor behind the emerging new "sciences of complexity" is the growing belief that "high-level" behavior of all complex systems can be traced back to essentially the same fundamental set of universal principles. Much of the study of complex systems consists of looking for the "low-level" underpinnings of universal patterns of "high-level" behavior.

Turning our attention now more to the subject of this study, the italicized question at the beginning of the previous paragraph can be rephrased as "What can we learn from how complex adaptive systems behave in general – as well as from the techniques and methodologies that have been developed for the theoretical analysis of real physical systems – that offers an alternative approach to land warfare modeling?"

The study of complex adaptive systems is predicated on the belief that while individual systems may differ in the details of their composition and internal dynamics, they nonetheless all share a general set of fundamental principles underlying their overall patterns of behavior. It is interesting to point out that one of the main reasons why complex adaptive systems are so attractive to study happens to also be one of the main reasons why it is very difficult to abstract a unifying theory for their behavior. Because complex adaptive systems typically involve
their behavior. Because complex adaptive systems typically involve nonlinear interactions, their overall behavior is usually more than just a simple sum of the behaviors of their parts. Nonetheless, while research in this still-developing field has yet to produce anything that comes close to being an all-encompassing "theory of complex systems," it has already uncovered many provocative and useful organizing principles.

General references include Bak [15], Cowan, et. al. [60], Holland [144], Kauffman [172], Lewin [196], Mainzer [205], Nadel [227]-[230], and Waldrop [311].

What Complex Systems Theory Is

Terms and concepts such as new sciences, chaos, complexity, complex systems theory, complex adaptive systems, and so on, are commonly used to denote the ostensibly same fundamental core of principles. While there is considerable overlap in meaning among the terms on this list, it is a mistake to believe that all of these terms can be used interchangeably.

Figure 2 shows a sketch of a loose but defensible definition of complex system theory (CST). The reader is urged to keep this image in mind throughout the ensuing discussion in this paper.

Figure 2. What is complexity?

They study of the behavior of collections of simple (and typically nonlinearly) interacting parts that can evolve and adapt to a changing environment

The simplest way to define complex systems theory is to contrast it with a conventional "classical physics" approach to the dynamics of simple systems. Ignoring the inherent complexities of individual problems, most of classical physics rests on the basic assumption that if only the initial state of a physical system – say $S_{\text{initial}}$ – is known, then the final state of the system – say, $S_{\text{final}}$ – can be obtained by a suitable function $f$. That is to say, classical physics essentially reduces to searching for a function $f$, such that $S_{\text{final}} = f(S_{\text{initial}})$. Classical physics also assumes that
both the initial and final states of the system – \( S_{\text{initial}} \) and \( S_{\text{final}} \), respectively – can be specified exactly by some finite set of variables.

If the physical system is simple enough, of course, these simplifying assumptions and description is entirely adequate. For example, if one holds an object at some height off the ground, the initial state of this "system" is completely specified by the parameters \((m, v_0=0, h)\), where \( m \) is the object's mass, \( v_0 \) is its initial velocity and \( h \) is its height from the ground. After the object is released, its final state, as it hits the ground, is easily found to be \((m, v = \sqrt{2gh}, h=0)\), where \( g \) is the gravitational constant \((g=9.8 \text{ m/s}^2)\).

In contrast, CST is concerned with more complicated systems, where "complicated" typically means that a system consists of a large number of mutually interrelated parts. In dealing with such systems, CST generalizes the conventional approach in two fundamental ways: (1) the final state, \( S_{\text{final}} \), is no longer assumed to be a function of the initial state alone, but can depend strongly on the path, \( P \), that the system follows in evolving from its initial to final states, and (2) the initial state is endowed with both an internal and external structure. CST can be described as the study of the behavior of collections of simple (and typically nonlinearly) interacting parts that can evolve.

Generally speaking, complex systems theory...

• is a general approach to understanding the overall behavior of a system composed of many nonlinearly interacting parts that is predicated on the premise that the system's behavior owes at least as much to how the system's parts all interact as to what those parts are

• teaches us that "complex behavior" is usually an emergent self-organized phenomenon built upon the aggregate behavior of very many nonlinearly interacting "simple" components.

• is an approach that tries to construct the minimal underlying rule set from which desired behaviors naturally emerge rather than hard-wiring in desired properties and/or behaviors from the start.

Complex adaptive system theory also assumes that systems are composed of interacting agents that continually adapt by changing their internal rules as the environment and their experience of that environment both evolve over time. Since a major component of an agent's environment consists of other agents, agents spend a great deal of their time adapting to the adaptation patterns of other agents.
What Complex Systems Theory Is Not

While it is important to understand what complex systems theory is, it is equally as important to understand what it is not. Complex systems theory is neither a "canned" algorithm that sits ready-and-able somewhere on a shelf, nor is it even a well-defined methodology:

- each problem must be approached on its own terms
- what is usually common to most approaches is what is borrowed from dynamical systems theory, computer science, information theory, biological/chemical pattern formation
- CST and CASs are best described as qualitative – not quantitative – interdisciplinary sciences; consequently, they are probably a poor modeling choice if numerical predictability is desired

<table>
<thead>
<tr>
<th>Table 2. Land combat as a complex adaptive system</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>nonlinear interaction</strong></td>
</tr>
<tr>
<td><strong>hierarchical structure</strong></td>
</tr>
<tr>
<td><strong>decentralized control</strong></td>
</tr>
<tr>
<td><strong>self-organization</strong></td>
</tr>
<tr>
<td><strong>nonequilibrium order</strong></td>
</tr>
<tr>
<td><strong>adaptation</strong></td>
</tr>
<tr>
<td><strong>collectivistic dynamics</strong></td>
</tr>
</tbody>
</table>

Land Combat as a Complex Adaptive System?

Military conflicts, particularly land combat, have almost all of the key features of complex adaptive systems (see table 2):

- **Nonlinear interaction**
  - friendly and enemy forces are composed of a large number of non-linearly interacting "parts"
○ combat is not just an aggregate of many smaller-scale conflicts, but is a complex system composed of parts whose action and pattern of behavior depends on the action and pattern of behavior of other (nearby and not-so-nearby) parts

○ "parts" interact in part...
  - *locally* – according to default combat "doctrine," to what "neighboring" parts are themselves doing, and to explicit orders issued by local commanders

    and, in part

  - *globally* – according to the orders issued by global commanders

○ sources of nonlinearity include:
  - feedback loops in C2 hierarchy
  - interpretation of, and adaptation to, enemy actions; i.e. nonlinear feedback among enemy combatants (measure \(\rightarrow\) *counter*-measure \(\rightarrow\) *counter*-countermeasure \(\rightarrow\) ...) and between combatants and environment
  - elements of chance ("fog of war")
  - decision making process (in which a sequence of not necessarily optimal or desired events is often put into motion because of other seemingly insignificant events)

○ combat timelines and partial outcomes are often determined by (often unanticipated) local effects (exponential divergence of trajectories)

• *Hierarchical structure*

  ○ parts are organized in a (command and control) hierarchy

• *Decentralized control*

  ○ despite the presence of "global commanders," who have a (global, albeit imprecise) view of the overall combat arena, there is no master "voice" that dictates the actions of each and every combatant

• *Self-organization*
local action, which often appears "chaotic," induces long-range order

command and control tends to organize what is otherwise disorganized action

**Nonequilibrium order**

- military conflicts, by their very nature, typically proceed far from equilibrium
- there is no unique "solution" – no stable state – towards which a battle evolves
- the lifeblood of complex adaptive systems is novelty and nonequilibrium. Military campaigns likewise depend on the creative leadership of their commanders, success or failure often hinging on the brilliant tactics conceived of in the heat of combat or the mediocre one that is issued in its place.

**Adaptation**

- their parts, in order to survive, must continually adapt to a changing combat environment (new strategies and tactics must be conceived of and implemented on-the-spot and in immediate response to changes in the environment)
- each combatant comes into a conflict armed with a set of default rules ("doctrine"), a goal (or goals) and hardware designed to facilitate the implementation of doctrine. The success or failure of a campaign depends on how well each combatant adapts to the continually changing combat environment, which includes the functioning and adaptation of both friendly and enemy combatants.
- actions and outcomes of actions are as much a function of the internal "human element" (reasoning capacity, unpredictability, inspiration, accident, etc.) as they are of the hardware

**Parts are more like "niches" than "parts"**

- their parts, particularly those represented by the lowest level combatant and as long as the warfighting skills of combatants exceed some threshold warfighting skill level, are essentially interchangeable
Collectivism

- there is continual feedback between the behavior of (low-level) combatants and the (high-level) C^2 hierarchy

The central thesis of this paper is that these largely conceptual connections between properties of land warfare and properties of complex systems in general can be extended to forge a set of practical connections as well. That is to say, that land warfare does not just look like a complex system on paper, but can be well characterized in practice using the same basic principles that are used for discovering and identifying behaviors in complex systems.

Redefined Conventions

Looking at land warfare through CAS-colored glasses naturally requires us to redefine the conventions by which military conflicts have traditionally been viewed:

- where conventional wisdom sees combat as essentially a head on collision between two massive (and perhaps slightly malleable) billiards, obeying a Newtonian-physics-like calculus-of-interaction, the CAS approach sees a self-organized hierarchy of evolving activity of two interacting fluids, in which global patterns of combat emerge out of an evolving substrate of low-level local interaction rules

- where conventional wisdom focuses on losses and attrition, the CAS approach highlights the evolving struggle to survive and emerging global patterns of (locally unanticipated) behavior

- where conventional wisdom asks "What are the consequences of this strategy?" the CAS approach tries instead to objectively map out the entire space of possible strategies.

Looking at land warfare through CAS-colored glasses also naturally alters the types of questions that are asked of "models of reality." The typical kinds of CAS questions one might ask of combat models include:

- What elements of combat are universal? That is, what elements of combat transcend the details of the individual components of which they are composed?

- What kinds of self-organized behaviors emerge out of a system obeying a well-defined combat-doctrine? To what extent is a military force
greater than the simple sum of its parts? To what extent does military doctrine enhance force synergy?

- What does the global "decision space" look like? That is, how do the various doctrinal, tactical and/or strategic elements all correlate with one another? Map out the entire possibility phase space of options available for all local and global combat units.
Nonlinear dynamics and Chaos

"Not only in research, but also in the everyday world of politics and economics, we would all be better off if more people realized that simple dynamical systems do not necessarily lead to simple dynamical behavior."

- R. M. May

So concludes Robert May in his well-known 1976 Nature review article [208] of what was then known about the behavior of first-order difference equations of the form \( x_{n+1} = F(x_n) \). What was articulated by a relatively few then, is now generally regarded as being the central philosophical tenet of chaos theory: complex behavior need not stem from a complex underlying dynamics.

Introduction

In this section we introduce the basic theory and concepts of nonlinear dynamics and chaos. The discussion begins with a definition of a deterministic dynamical system, which for our purposes we define simply as any physical system for which there exists a well-defined prescription, either in terms of differential or difference equations, for calculating the future behavior given only the system’s initial state. Given that such systems evolve deterministically in time, one might reasonably expect them to behave regularly at all times. After all, each successive state is a uniquely prescribed function of the preceding state. Chaos theory shows, however, that this naïve intuition is wrong, and that perfectly well-defined, deterministic, but nonlinear dynamics, often leads to erratic and apparently random motion. Moreover, the dynamics itself need not be at all complicated.

Short History

Table 3 shows a brief chronology of some of the milestone events in the study of nonlinear dynamics and chaos.

Chaos was arguably born, at least in concept, at the turn of the last century with Henri Poincare’s discovery in 1892 that certain orbits of three or more interacting celestial bodies can exhibit unstable and unpredictable behavior. A full proof that Poincare’s unstable orbits are chaotic, due to Smale, appeared only 70 years later. E. N. Lorenz’ well-known paper in which he showed that a simple set of three coupled, first order, nonlinear differential equations describing a simplified model of the atmosphere can lead to completely chaotic trajectories was published a year after Smale’s proof, in 1963. As in Poincare’s case, the general significance of Lorenz’ paper was not appreciated until many years after its publication. The formal rigorous study of deterministic chaos began in earnest with Mitchell
Feigenbaum's discovery in 1978 of the universal properties in the way nonlinear dynamical systems approach chaos.

Table 3. Some historical developments in the study of nonlinear dynamics and chaos

<table>
<thead>
<tr>
<th>Year</th>
<th>Researchers</th>
<th>Discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>1875</td>
<td>Weierstrass</td>
<td>constructed everywhere continuous and nowhere differentiable function</td>
</tr>
<tr>
<td>1890</td>
<td>King Oscar II of Sweden</td>
<td>offered prize for 1st person to solve the n-body problem to determine the orbits on n-celestial bodies and thus prove the stability of the solar system; this problem remains unsolved in 1995</td>
</tr>
<tr>
<td>1892</td>
<td>Poincare</td>
<td>in the course of studying celestial motion, discovered that the (&quot;homoclinic&quot;) orbit of three or more interacting bodies can exhibit unstable and unpredictable behavior (chaos is born!)</td>
</tr>
<tr>
<td>1932</td>
<td>Birkhoff</td>
<td>observed what he called &quot;remarkable curves&quot; in the dynamics of the plane with itself</td>
</tr>
<tr>
<td>1954</td>
<td>Kolmogorov</td>
<td>discovered that motion in phase space of classical mechanics is neither completely regular nor completely irregular, but that trajectory depends on the initial conditions; KAM theorem</td>
</tr>
<tr>
<td>1962</td>
<td>Smale</td>
<td>mathematical proof that Poincare's homoclinic orbits are chaotic</td>
</tr>
<tr>
<td>1963</td>
<td>Lorenz</td>
<td>first systematic analysis of chaotic attractors in simplified model of atmospheric air currents; coined the &quot;Butterfly effect&quot;</td>
</tr>
<tr>
<td>1970</td>
<td>Mandelbrot</td>
<td>coined the term &quot;fractal&quot; and suggested applicability to a wide variety of natural phenomena</td>
</tr>
<tr>
<td>1971</td>
<td>Ruelle, Takens</td>
<td>suggest new mechanism for turbulence: strange attractors</td>
</tr>
<tr>
<td>1975</td>
<td>Li, Yorke</td>
<td>use &quot;chaos&quot; to denote random output of deterministic mappings</td>
</tr>
<tr>
<td>1976</td>
<td>May</td>
<td>wrote important review article in Nature on complicated dynamics of population dynamics models</td>
</tr>
<tr>
<td>1978</td>
<td>Feigenbaum</td>
<td>discovered universal properties in the way nonlinear systems approach chaos</td>
</tr>
<tr>
<td>1990</td>
<td>Ott, Grebogi, Yorke</td>
<td>beginning of chaos control theory</td>
</tr>
<tr>
<td>1990</td>
<td>Pecora</td>
<td>beginning of synchronization of chaotic systems</td>
</tr>
</tbody>
</table>

The term "chaos" was first coined by Li and Yorke in 1975 to denote random output of deterministic mappings. More recently, in 1990, Ott, Grebogi and Yorke suggested that certain properties of chaotic systems can be exploited to control chaos; that is, to redirect the
chaotic system into another desired orbit. Ironically, chaos can be "controlled" precisely because of its inherent instabilities, and there is no counterpart "control theory" for nonchaotic systems (see page 59).

Experimentally, deterministic chaos has by now been observed in just about every conceivable physical system that harbors some embedded nonlinearity: arms races, biological models for population dynamics, chemical reactions, fluids near the onset of turbulence, heart beat rhythms, josephson junctions, lasers, neural networks, nonlinear optical devices, planetary orbits, etc.

Fractals – that is, self-similar objects that harbor an effectively infinite number of layers of detail – were (formally) born in 1875, when the mathematician Weierstrass had constructed an everywhere continuous but nowhere differentiable function, though Weierstrass neither coined the term nor was, in his time, able to fully appreciate the complexity of his own creation. A fuller understanding of fractals had to await the arrival of the speed and graphics capability of the modern computer. The term "fractal" was introduced by Mandelbrot about a hundred years after Weierstrass' original construction.

**Dynamical Systems**

A dynamical system – as it is typically understood by physicists – is any physical system that evolves in time according to some well-defined rule. Its state is completely defined at all times by the values of N variables, \( x_1(t), x_2(t), \ldots, x_N(t) \), where \( x_i(t) \) represent any physical quantity of interest (position, velocity, temperature, etc.). The abstract space in which these variables "live" is called the phase space \( \Gamma \). Its temporal evolution is specified by an autonomous system of \( N \), possibly coupled, ordinary first-order differential equations:

\[
\frac{dx_1}{dt} = F_1(x_1, x_2, \ldots, x_N; \alpha_1, \alpha_2, \ldots, \alpha_M),
\]

\[
\frac{dx_2}{dt} = F_2(x_1, x_2, \ldots, x_N; \alpha_1, \alpha_2, \ldots, \alpha_M),
\]

\[\vdots\]

\[
\frac{dx_N}{dt} = F_N(x_1, x_2, \ldots, x_N; \alpha_1, \alpha_2, \ldots, \alpha_M),
\]

\footnote{Note that nonlinearity is a necessary, but not sufficient condition for deterministic chaos. Linear differential of difference equations can be solved exactly and do not lead to chaos.}
where $\alpha_1, \alpha_2, ..., \alpha_M$ are a set of M control parameters, representing any external parameters by which the evolution may be modified or driven. The temporal evolution of a point $x(t) = (x_1(t), x_2(t), ..., x_N(t))$ traces out a trajectory (or orbit) of the system in $\Gamma$. The system is said to be linear or nonlinear depending on whether $F = (F_1, F_2, ..., F_N)$ is linear or nonlinear. Nonlinear systems generally have no explicit solutions.

Once the initial state $x(t=0)$ of the system is specified, future states, $x(t)$, are uniquely defined for all times $t$. Moreover, the uniqueness theorem of the solutions of ordinary differential equations guarantees that trajectories originating from different initial points never intersect.

In studying deterministic chaos, one must make a distinction between chaos in dissipative systems (such as a forced pendulum with friction) and conservative systems (such as planetary motion); see below.

### Discrete-time Poincaré maps

A convenient method for visualizing continuous trajectories is to construct an equivalent discrete-time mapping by a periodic "stroboscopic" sampling of points along an orbit. One way of accomplishing this is by the so-called Poincaré map (or surface-of-section) method. In general, an $(N-1)$-dimensional surface-of-section $S$ in the phase space $\Gamma$ is chosen, and we consider the sequence of successive intersections $I_1, I_2, ..., I_i, ...$ of the flow $x(t)$ with $S$. Introducing a system of coordinates, $y_1, y_2, ..., y_{N-1}$, on $S$ and representing the intersections $I_i$ by coordinates $y_{i,1}, y_{i,2}, ..., y_{i,N-1}$, the system of differential equations is replaced by the discrete-time Poincare mapping (see figure S):

$$y'_i = G_1(y_{i,1}, y_{i,2}, ..., y_{i,N-1}; \alpha_1, \alpha_2, ..., \alpha_M),$$

$$y'_{i+1} = G_2(y_{i,1}, y_{i,2}, ..., y_{i,N-1}; \alpha_1, \alpha_2, ..., \alpha_M),$$

$$...$$

$$y'_{i,N-1} = G_{N-1}(y_{i,1}, y_{i,2}, ..., y_{i,N-1}; \alpha_1, \alpha_2, ..., \alpha_M).$$

3 If $f$ is a nonlinear function or an operator, and $x$ is a system input (either a function or variable), then the effect of adding two inputs, $x_1$ and $x_2$, first and then operating on their sum is, in general, not equivalent to operating on two inputs separately and then adding the outputs together; i.e. $f(x+y)$ is, in general, not equal to $f(x) + f(y)$. 

24
Phase Space Volumes

Consider a small rectangular volume element $\Delta V$ around the point $x_0$. For discrete-time Poincare maps of the form $x_{n+1} = G(x_n)$, the rate of change of $\Delta V$ — say, $\Lambda$ — is given by the absolute value of the Jacobian of $G$:

$$\Lambda = |J| = \left| \det \left( \frac{\partial G}{\partial x} \right) \right|.$$  

Since the motion in phase space is typically bounded, we know that volumes do not, on average, expand; i.e. $\Lambda$, and therefore the Jacobian $J$, are not positive. On the other hand, the behavior of systems for which $\Lambda < 0$ (called **dissipative systems**) is very different from the behavior of systems that have $\Lambda = 0$ (called **conservative systems**).

Dissipative Dynamical Systems

Dissipative systems — whether described as continuous flows or Poincare maps — are characterized by the presence of some sort of "internal friction" that tends to contract phase space volume elements. Contraction in phase space allows such systems to approach a subset of the phase space called an **attractor**, $A \subset \Gamma$, as $t \to \infty$. Although there is no universally accepted definition of an attractor, it is intuitively reasonable to demand that it satisfy the following three properties:

- **Invariance**: $A$ is invariant under the map $F$ — i.e. $FA = A$
• *Attraction*: there is an open neighborhood $B$ containing $A$ such that all points $x(t)$ in $B$ approach $A$ as $t \to \infty$; the set of initial points $x_0(t=0)$ such that $x(t)$ approaches $A$ is called the *basin of attraction* of $A$.

• *Irreducibility*: $A$ cannot be partitioned into two nonoverlapping invariant and attracting pieces; a more technical demand is that of topological transitivity – there must exist a point $x^*$ in $A$ such that for all $x$ in $A$ there exists a positive time $T$ such that $x^*(T)$ is arbitrarily close to $x$.

The simplest possible attractor is a *fixed point*, for which all trajectories starting from the appropriate basin-of-attraction eventually converge onto a single point. For linear dissipative dynamical systems, fixed point attractors are in fact the only possible type of attractor. Nonlinear systems, on the other hand, harbor a much richer spectrum of attractor-types. For example, in addition to fixed-points, there may exist periodic attractors such as *limit cycles* for two-dimensional flows or doubly periodic orbits for three-dimensional flows. There is also an intriguing class of attractors that have a very complicated geometric structure called *strange attractors*.

**Strange Attractors**

The motion on strange attractors exhibits many of the properties normally associated with completely random or chaotic behavior, despite being well-defined at all times and fully deterministic. More formally, a strange attractor $A_\delta$ is an attractor (meaning that it satisfies the three properties of attractors given above) that also displays *sensitivity to initial conditions*. In the case of a one-dimensional map, $x_{n+1} = f(x_n)$, for example, this means that there exists a $\delta > 0$ such that for all $x$ in $A_\delta$ and any open neighborhood $U$ of $x$, there exists $x^*$ in $U$ such that $|f^n(x) - f^n(x^*)| > \delta$. The basic idea is that initially close points become exponentially separated for sufficiently long times. This has the important consequence that while the behavior of each initial point may be accurately followed for short times, prediction of long time behavior of trajectories lying on strange attractors becomes effectively impossible. Strange attractors also frequently exhibit a self-similar or fractal structure.

**Deterministic Chaos**

"Chaos is a name for any order that produces confusion in our minds."

- G. Santayana

What is deterministic chaos? Despite the over three decades of research and countless books and papers that have been written on the
subject of deterministic chaos, there is still no generally accepted
definition. Intuitively, deterministic chaos is the irregular or random
appearing motion in nonlinear dynamical systems whose dynamical
laws uniquely determine the time evolution of the state of the system
from a knowledge of its past history. It is not due to either external
noise, the system having an infinite number of degrees-of-freedom or
any quantum-mechanical uncertainty. The source of the observed
irregularity in deterministic chaos is an intrinsic sensitivity to initial
conditions.

A more mathematically rigorous definition of chaos, that holds for
both continuous and discrete systems, is due to Devaney [72]. Let V be
a set. A map \( f: V \rightarrow V \) is said to be chaotic on \( V \) if (1) \( f \) has sensitive
dependence on initial conditions, (2) \( f \) is topologically transitive, and
(3) periodic points are dense in \( V \). Devaney states:

"To summarize, a chaotic map must possesses three
ingredients: unpredictability, indecomposability, and the
element of regularity. A chaotic system is unpredictable
because of sensitive dependence on initial conditions. It
cannot be broken down or decomposed into two subsystems
(two invariant open subsets) which do not interact under \( f \)
because of topological transitivity. And, in the midst of this
random behavior, we nevertheless have an element of
regularity, namely the periodic points which are dense."  

Several examples of deterministic chaos are discussed below.

**Conservative Dynamical Systems**

In contrast to dissipative dynamical systems, conservative systems
preserve phase-space volumes and hence cannot display any attracting
regions in phase space. Consequently, there can be no fixed points, no
limit cycles and no strange attractors. However, there can still be
chaotic motion in the sense that points along particular trajectories
may show sensitivity to initial conditions. A familiar example of a
conservative system from classical mechanics is that of a Hamiltonian
system. Although the chaos exhibited by conservative systems often
involves fractal-like phase-structures, the fractal character is of an
altogether different kind from that arising in dissipative systems.

---

4. A topologically transitive orbit is an orbit such that, for all pairs of
regions in the phase space, the orbit at some point visits each region of the
pair. That is to say, it is always possible to eventually get from one area around
a state to an area around any other area by following the orbit.

5. A set of points \( X \) is dense in another set \( Y \) if an arbitrarily small area
around any point in \( Y \) contains a point in \( X \).
Example #1: The Bernoulli Shift Map

Despite bearing no direct relation to any physical dynamical system, the one-dimensional discrete-time piecewise linear Bernoulli Shift map nonetheless displays many of the key mechanisms leading to deterministic chaos. The map is defined by

\[ x_{n+1} = f(x_n) = 2x_n \mod 1, \quad 0 < x_0 < 1, \]

where \( x \mod 1 = x - \text{Integer}(x) \), and \( \text{Integer}(x) \) is the "integer part of" \( x \).

We are interested in the properties of the sequence of values \( x_n \), \( x_n = f(x_{n-1}) = f^2(x_{n-2}) = \ldots \) or the orbit of \( x_n \) generated by successive applications of the Bernoulli shift to the initial point, \( x_0 \).

In turns out that the most convenient representation for the initial point, \( x_0 \), is as a binary decimal. That is, we write

\[ x_0 = \sum \frac{\alpha_i}{2^i} = \frac{\alpha_1}{2} + \frac{\alpha_2}{4} + \ldots = 0.\alpha_1\alpha_2\alpha_3\ldots, \]

where \( \alpha_i \) is equal to either 0 or 1 for all \( i \).

For example, the binary expansion of \( 1/3 = 0/2 + 1/2^2 + 0/2^3 + 1/2^4 + \ldots = 0.0101 \), where 01 means that the sequence "01" is repeated ad-infinimtum. Expansions for arbitrary rationals \( r = p/q \), where \( p \) and \( q \) are integers, are relatively easy to calculate. Expansions for irrational numbers may be obtained by first finding a suitably close rational approximation. For example, \( \pi - 3 \approx 4703/33215 = 0.001001000011 \), which is correct to 12 binary decimal places.

This binary decimal representation of \( x_0 \) makes clear why this map is named the Bernoulli "shift." If \( x_0 < 1/2 \), then \( \alpha_i = 0 \); if \( x_0 > 1/2 \), then \( \alpha_i = 1 \). Thus

\[ x_1 = f(x_0) = \begin{cases} 2x_0 & \text{if } \alpha_1 = 0 \\ 2x_0 - 1 & \text{if } \alpha_1 = 1 \end{cases} \Rightarrow f(x_0) = 0.\alpha_2\alpha_3\alpha_4\ldots \]

In other words, a single application of the map \( f \) to the point \( x_0 \) discards the first digit and shifts to the left all of the remaining digits in the binary decimal expansion of \( x_0 \). In this way, the nth iterate is given by \( x_n = \alpha_{n+1} \alpha_{n+2} \alpha_{n+3} \ldots \)

**What are the properties of the actual orbit of \( x_0 \)?** Since \( f \) effectively reads off the digits in the binary expansion of \( x_0 \), the properties of the orbit depend on whether \( x_0 \) is rational or irrational. For rational \( x_0 \), orbits are both periodic and dense in the unit interval; for irrational \( x_0 \), orbits are nonperiodic, with the attractor being equal to the entire unit.
interval. Moreover, the Bernoulli shift is ergodic. That is to say, because any finite sequence of digits appears infinitely many times within the binary decimal representation of almost all irrational numbers in [0,1] (except for a set of measure zero), the orbit of almost all irrationals approaches any x in the unit interval to within an arbitrarily small distance an infinite number of times.

We now use the Bernoulli shift to illustrate four fundamental concepts that play an important role in deterministic chaos theory:

- stability
- predictability
- deterministic randomness
- computability

**Stability**

Chaotic attractors may be distinguished from regular, or nonchaotic, attractors by being unstable with respect to small perturbations to the initial conditions; this property is frequently referred to as simply a sensitivity to initial conditions. However, while all Bernoulli shift orbits are generally unstable in this sense, only those originating from irrational x₀ are chaotic. Suppose that two points, x₀ and x₀', differ only in the nth place of their respective binary decimal expansions. By the nth iterate, the difference between their evolved values, |fⁿ(x₀) - fⁿ(x₀')|₁, will be expressed in the first digit; i.e., arbitrarily small initial differences – or "errors" – are exponentially magnified in time. If |x₀ - x₀'|₁ ~ 10⁻¹, for example, their respective orbits would differ by order ~1 by the 100th iterate. Physically, we know that any measurement will have an arbitrarily small, but inevitably finite, error associated with it. In systematically magnifying these errors, nonlinear maps such as the Bernoulli shift effectively transform the information originating on microscopic length scales to a form that is macroscopically observable.

**Predictability**

Exponential divergence of orbits places a severe restriction on the predictability of the system. If the initial point x₀ is known only to within an error δx₀, for example, we know that this error will grow to δxₙ = exp(n ln 2) δx₀ (mod 1) by the nth iteration. The relaxation time, t, to a statistical equilibrium – which is defined as the number of iterations required before we reach a state of total ignorance as to the
minimum $n$ such that $\delta x_n \sim 1$ is therefore given by $\tau \sim \ln|\delta x_0|/\ln 2$. For all times $t > \tau$, the initial and final states of the system will be causally disconnected.

**Deterministic Randomness**

On the one hand, the Bernoulli shift is a linear difference equation that can be trivially solved for each initial point $x_0$: $x_n = 2^n x_0 \pmod{1}$. Once an initial point is chosen, the future iterates are determined uniquely. As such, this simple system is an intrinsically deterministic one. On the other hand, look again at the binary decimal expansion of a randomly selected $x_0$. This expansion can also be thought of as a particular semi-infinite sequence of coin tosses,

$$x_0 = 0.011010011010001... \Leftrightarrow x_0 = 0.THHTHTTHHHHT...$$

in which each 1 represents heads and each 0 represents tails. In this way, the set of all binary decimal expansions of $0 < x_0 < 1$ can be seen as being identical to the set of all possible sequences of random coin tosses. Put another way, if we are merely reading off a string of digits coming out of some "black-box," there is no way of telling whether this black-box is generating the outcome by flipping a non-biased coin or is in fact implementing the Bernoulli shift for some precisely known initial point. Arbitrarily selected $x_0$ will therefore generate, in a strictly deterministic manner, a random sequence of iterates, $x_0, x_1, x_2, ...$ Notice, however, that the Bernoulli shift generated "randomness" is of an altogether different character from that exhibited by the temporal sequence of center-site values in some cellular automaton systems (see page 81). While a cellular automaton system generates random sequences from manifestly nonrandom simple initial seeds, the Bernoulli shift effectively unravels the randomness that is already present in the initial state. Moreover, it is important to point out that while one is always assured of randomly selecting an irrational $x_0$ (with probability one) – by virtue of the fact that rationals only occupy a space of measure zero – one is at the same time limited in a practical computational sense to working with finite, and therefore rational, approximations of $x_0$. The consequences of this fact are discussed below.

**Computability**

While one can formally represent an arbitrary point $x$ by the infinite binary-decimal expansion $x = 0.\alpha_1\alpha_2\alpha_3...$, in practice one works only with the finite expansion, $x = 0.\alpha_1\alpha_2\alpha_3...\alpha_n$. Conversely, any sequence of coin tossings is also necessarily finite in duration and therefore defines only a rational number.
Given this restriction, in what sense are chaotic orbits computable? When implemented on a computer, for example, a single iteration of the Bernoulli map is realized by a left shift of one bit followed by an insertion of a zero as the rightmost bit. Since any $x_n$ is stored as a finite-bit computer word, the result is that all $x_n$ are eventually mapped to the (stable fixed point) $x_n^*$:

\[
\begin{align*}
x_0 &= 0 \quad \alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \ldots \quad \alpha_{n-1} \quad \alpha_n \\
x_1 &= 0 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4 \quad \ldots \quad \alpha_n \quad 0 \\
\vdots & \quad \quad \vdots \\
x_n &= 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0
\end{align*}
\]

All of the points of a finite length orbit, $x_0, x_1, \ldots, x_m$, may therefore be assured of having at least $m$-bit accuracy by computing $x_n$ to $n+m$ bits. A number $x$ is said to be computable if its expansion coefficients $\alpha_i$ may be algorithmically generated to arbitrarily high order. Thus, so long as the initial point $x_0$ is itself a computable irrational number, its orbit will be chaotic and computable. One can show, however, that there are many more noncomputable irrationals than computable ones.

**Example #2: The Logistic Map**

Just as the Bernoulli shift map provides important insights into some of the fundamental properties of dynamical chaos, the logistic map is arguably the simplest (continuous and differentiable) nonlinear system that captures most of the key mechanisms responsible for producing dynamical chaos. Indeed, the logistic map appears to capture much of the essence of a whole class of real-world phenomena, including that of the transition to turbulence in fluid flows.

Although the basic properties of the logistic map have been studied for at least forty years, the most profound revelations are due to Mitchell Feigenbaum’s analyses in the mid 1970s, culminating in his universality theory. Feigenbaum observed that the "route to chaos" as found in the logistic map in fact occurs (apart from a few mild technical restrictions) in all first-order difference equations $x_{n+1} = f(x_n)$, where the function $f(x)$ (after a suitable rescaling) has a single maximum on the unit interval $[0,1]$. Moreover, the transition to chaos is characterized by a scaling behavior governed by universal constants whose value depends only on the order of the maximum of $f(x)$. Because the properties of the logistic map underlie so much of what generally goes under the rubric of "chaos theory," we provide a short overview of the behavior of this simple dynamical system.
It is ironic that such an intensely computational mathematical science as chaos theory owes much of its modern origin to calculations that were performed not on a large mainframe computer but on a simple programmable pocket calculator, a Hewlett-Packard HP-65. Feigenbaum likes to point out that had he not had time to observe each and every step of the evolution of the logistic map, it is unlikely that he would have been able to see enough of the embedded “patterns” from which he ultimately induced his universality theory. An important lesson to be taken away from this is that minimal modeling does not necessarily lead only to trivial observations.

Figure 4. $x_n$ versus $n$ for the logistic map using four different values of $\alpha$.

![Graphs showing the logistic map for different values of $\alpha$.](image)

**Definition**

The logistic map is a one-dimensional nonlinear discrete difference equation with a single control parameter, $\alpha$:

$$x_{n+1} = f(x_n) = \alpha x_n (1 - x_n), \quad 0 \leq x_0 \leq 1, \quad 0 \leq \alpha \leq 4.$$
As long as the single control parameter, $\alpha$, is positive and less than or equal to four, the orbit of any point $x_0$ remains bounded on the unit interval. Notice that there are two antecedents, $x_n$ and $x'_n$, for each point $x_{n+1}$, so that, like the Bernoulli map, this map is also noninvertible.

Now consider the behavior of the orbits of the logistic map as a function of the parameter $\alpha$. Figure 4 illustrates the fact that the behavior of this map is strongly dependent on the value of $\alpha$.

**Fixed Point Solutions**

We begin by asking whether there are any values of $\alpha$ for which the system has fixed points. Solving the fixed-point equation

$$x^* = f(x^*) = \alpha x^*(1 - x^*),$$

we find two such points: $x^* = x^*_{(0)} = 0$ and $x^* = x^*_{(1)} = (\alpha - 1)/\alpha$. In order for $x^*_{(1)}$ to be in the unit interval, we must have that $\alpha \geq 1$. What of the stability of these two points? As we have already seen in the case of the Bernoulli map, the divergence of initially close by points is a crucial issue in the analysis of the dynamical behavior.

Given a fixed point, $x^*$, the subsequent evolution of a nearby point, $x^{*'} = x^* + \epsilon$, where $\epsilon \ll 1$, may be determined by substituting $x_n = x^* + \epsilon_n$ and $x_{n+1} = x^* + \epsilon_{n+1}$ into the fixed-point equation $f(x^*) = x^*$ and leaving only the terms that are linear in $\epsilon_n$ and $\epsilon_{n+1}$:

$$\epsilon_{n+1} = \alpha(1 - 2x^*)\epsilon_n + O(\epsilon_n^2),$$

where $O(\epsilon_n^2)$ represents terms of order $\epsilon_n^2$ and smaller. We find that, regardless of the initial point $x_0$, the deviations from the fixed point $x^*_{(0)}$ decrease exponentially fast for all $\alpha < 1$. That is to say, all points $x_0 \in [0, 1]$ are attracted to the fixed point $x^* = 0$.

At the critical value $\alpha = \alpha_c = 1$, $x^*_{(0)}$ becomes unstable and the $\alpha$-dependent fixed point $x^*_{(1)}$ becomes stable. This exchange of stability between two fixed points of a map is known as a transcritical bifurcation. By using the same linear-stability analysis as above, we see that $x^*_{(1)}$ remains stable if $-1 < \alpha(1 - x^*_{(1)}) < 1$, or for all $\alpha$ such that $1 < \alpha < 3$.

For $\alpha > 3$, neither $x^* = 0$ nor $x^* = 1 - \alpha^{-1}$ is stable. Instead, the stable orbit is a period-2 limit cycle consisting of two points – $\{x_1^*, x_2^*\}$ – where $x_1^*$ and $x_2^*$ are each fixed points of the doubly iterated map $f^2(x) = \alpha x^*(1 - x^*)$. 33
f[f(x)]. At $\alpha + \sqrt{6} \approx 3.44949$ the period-2 attractor loses stability and is replaced by a stable period-4 orbit.

As $\alpha$ increases still further, the system undergoes an infinite sequence of successive period doubling "bifurcations":

- a stable period-$2^n$ orbit exists for all $\alpha$ such that $\alpha_{n-1} < \alpha < \alpha_n$

- at the nth critical value of $\alpha$ – i.e. at $\alpha_n$ – all points of the $2^{n-1}$ cycle simultaneously become unstable, and the system becomes attracted to a new stable period-$2^n$ cycle for $\alpha_n < \alpha < \alpha_{n+1}$

- While the period of the limit-cycles approaches infinity as $n \to \infty$, the distance between successive critical $\alpha$’s rapidly decreases: $\alpha_1 = 3.0$, $\alpha_2 = 3.44949...$, $\alpha_3 = 3.54409...$, $\alpha_4 = 3.56441...$, $\alpha_5 = 3.56876...$, ..., $\alpha_\infty = 3.5699456...$

Figure 5. Schematic representation of first few bifurcations in the logistic map

---

**Universality**

Feigenbaum’s important discovery consisted of the following two quantitative observations (see figure 5):

1. **Critical Parameter Convergence**: Feigenbaum found that the convergence rate of the critical parameters, $\alpha_n$, is geometric; i.e. that
\[ \alpha_n \text{ scale as } \alpha_n = \alpha_\infty - c\delta^{-n}, \text{ where } N \gg 1 \text{ and } 'c' \text{ and } '\delta' \text{ are constants.} \]

In fact,

\[ \delta = \lim_{n \to \infty} \frac{\alpha_n - \alpha_{n-1}}{\alpha_{n-1} - \alpha_n} = 4.6692016091... \]

This value of \( \delta \), known as the Feigenbaum constant, is the same for all one-dimensional maps \( f(x) \), where \( f(x) \) has a single quadratic maximum on the unit interval. For example, while the absolute values of the set of critical \( \alpha \)'s – \( \alpha_1, \alpha_2, ..., \alpha_n, ... \) – as calculated for the system \( x_{n+1} = \alpha' \sin(\pi x_n) \) will be different from the set of \( \alpha \)'s computed for the logistic map, their geometric convergence proceeds according to the same rate \( \delta \).

2. **Scaling of Branch Splittings:** Define a supercycle to be a cycle, \( \{x^*_n\} \), such that \( \prod_{i=1}^{2^n} f'(x^*_i) = 0 \). In particular, since \( f'_\alpha(x) = 0 \) only when \( x = 1/2 \), we know that the point \( x = 1/2 \) must always be an element of a supercycle. Define \( d_n \) to be the smallest distance between \( x = 1/2 \) and the nearest other point on the same \( 2^n \)-supercycle. Let \( \bar{\alpha}_n \) define the new nth critical value of \( \alpha \), \( \alpha_n < \bar{\alpha}_n < \alpha_{n+1} \) at which \( x = 1/2 \) becomes an element of the \( 2n \)-cycle.

Feigenbaum's second observation was that the relative scale of successive splittings of \( \bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3, ... \) approaches another universal constant, \( \Delta \):

\[ \Delta = \lim_{n \to \infty} \frac{d_n}{d_{n-1}} = 2.5029078750... \]

The convergence rate of \( \bar{\alpha} \) follows the same form as for \( \alpha_n \):

\[ \bar{\alpha}_n - \bar{\alpha}_\infty = c'\delta^{-n}, \text{where } \delta \text{ is the same as above and } \bar{\alpha}_\infty = \alpha_\infty. \]

**Behavior for } \alpha > \alpha_\infty**

What happens for \( \alpha > \alpha_\infty \)? An overview is provided in figure 6, which shows the numerically determined attractor sets for all \( \alpha \) in the range 2.9 - 4.0. Note the insert in the lower left corner, which shows a blowup view of the windowed region within the broad white band in the main figure.
Figure 6. Bifurcation plot for the logistic map

The general behavior of the logistic map for $\alpha > \alpha_\infty$ is summarized as follows:

- the attracting sets for many — but not all — $\alpha > \alpha_\infty$ are aperiodic and chaotic on various intervals of the unit interval $[0,1]$.

- As $\alpha$ increases, the chaotic intervals merge together by inverse bifurcations obeying the same $\delta$ and $\Delta$ scalings as in the $\alpha < \alpha_\infty$ region, until the attracting set becomes distributed over the entire unit interval at $\alpha = 4$.

- There are a large number of "windows" of finite width within which the attracting set reverts back to being a stable periodic cycle. Within these windows chaotic and periodic regions are densely interwoven. The largest such window — a snapshot of which is shown in the smaller boxed region in the lower left of figure 6 — corresponds to a stable period-3 cycle and spans the width $3.8284 < \alpha < 3.8415$.

- The periodic windows also harbor period-$m$ cycles, where $m = 3, 5, 6, \ldots$ that undergo period doubling bifurcations, $m \rightarrow 2m \rightarrow 4m \rightarrow \ldots$ at a set of critical parameters $\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots$ that again scale as
\[ \hat{\alpha}_n = \hat{\alpha}_\infty - \hat{\delta}^{-n} \], with the same universal \( \delta \) as in the \( \alpha < \alpha_\infty \) region.

- Other periodic windows harbor period triplings, quadruplings, etc., occurring at different sets of \( \{\hat{\alpha}_i\} \), but all of which scale in the familiar fashion (albeit with different universal constants \( \hat{\delta} \neq \delta \)).

Two-Dimensional Strange Attractors

"I have not spoken of the aesthetic appeal of strange attractors. These systems of curves, these clouds of points suggest sometimes fireworks or galaxies, sometimes strange and disquieting vegetal proliferations. A realm lies here to explore and harmonies to discover."  
— D. Ruelle

David Ruelle concludes his 1980 *The Mathematical Intelligencer* [270] review of strange attractors with this eloquent passage on the often strikingly beautiful patterns weaved by strange attractors.

Henon Map

One the simplest two-dimensional systems is an analogue of the logistic equation, introduced by Henon in 1976. It is defined by the equations

\[
\begin{align*}
x_{n+1} &= 1 - \alpha x_n^2 + y_n, \\
y_{n+1} &= \beta x_n,
\end{align*}
\]

where \( \alpha \) and \( \beta \) are constants; \( \alpha \) controls the extent of the nonlinearity, while \( \beta \) controls the degree of dissipation. Note that, unlike the logistic map, the Henon map is invertible. While noninvertibility is necessary for chaos in one-dimensional maps, it is not required in higher-dimensions.

Generally, the sequence of points, \((x_0,y_0), (x_1,y_1), ..., (x_n,y_n), ...\) either diverges to infinity (for \( x_0 \) large) or settles onto an attractor (for \((x_0,y_0)\) near the origin). A fixed-point analysis similar to the one performed earlier for the logistic equation may be carried out here to determine the behavior of the map as a function of \( \alpha \) and \( \beta \). For example, the fixed points are easily found to be \( x_* = \frac{1}{2\alpha} \left\{ -(1-\beta) \pm \sqrt{(1-\beta)^2 + 4\alpha} \right\} \), and \( y_* = \beta x_* \), where the point \((x_*, \beta x_*)\) is always unstable and \((x_*, \beta x_*)\) becomes unstable for \( \alpha > \frac{3}{4}(1-\beta)^2 \). For \( \beta = 3/10 \), both fixed points become unstable for \( \alpha = \alpha_1 = 0.3675 \) and a two-cycle is born. At \( \alpha = \alpha_2 = 0.9125 \) the two-cycle attractor becomes unstable and a four-cycle is
born. As \( \alpha \) is increased further, the system undergoes period-bifurcations obeying the same Feigenbaum constants \( \delta \) and \( \Delta \) that describe the period doublings in the logistic equation.

Figure 7 shows four snapshot views of the structure of the Henon strange attractor for \( \alpha = 1.4 \) and \( \beta = 0.3 \). In the figure, the second, third and fourth plots (counted clockwise starting from the upper left plot) provide enlargements of the small window regions shown in the immediately preceding plots. The attractor possesses two noteworthy properties:

Figure 7. The Henon Attractor (upper left-hand-side). Each succeeding image is an enlargement of the boxed segment of the immediately preceding image.

1. Exponential divergence of nearby trajectories: Just as global space-time CA patterns (see page 75) must be allowed to slowly emerge over many iteration steps before becoming recognizable and have a characteristic appearance that is often difficult if not impossible to predict beforehand from a knowledge of the CA rule alone, so too do strange attractors slowly, and in a highly irregular manner, form on a computer screen. While individual points jump around in an apparently haphazard fashion, it is only after several hundred have been plotted that the outline of the underlying attractor becomes clear. Numerical experiments confirm that the dynamics on this attractor are indeed chaotic.
2. Self-Similarity: What is most obvious from comparing the sequence of magnifications of an isolated section of the Henon attractor as shown figure 7, is its transverse self-similarity, or its Cantor-set-like structure. Since the map is area-contracting, the attracting set must have zero area (just as the classical Cantor-set has zero length); its Hausdorff dimension is nonintegral, however, which is a characteristic feature of fractals, see page 50). Moreover, although this may not be clear from the figure, the Henon attractor has an inhomogeneous structure. That is to say, the probabilities, $p_i$, for a point of the attractor to be in the $i^{th}$ band in each of the "blown-up" images, are all different.

Qualitative Characterization of Chaos

What are the criteria by which dynamical systems can be judged to be chaotic? Suppose you are given a dynamical system, $S$, or a set of time-series data of $S$'s behavior of the form

$$\zeta(t) = \zeta(t_0), \zeta(t_1), \zeta(t_2), ..., \zeta(t_N),$$

where $\zeta(t_i)$ represents the state of $S$ at time $t_i$ and $S$'s state is sampled every $t_{i+1} = t_i + \Delta t$ time steps for some fixed $\Delta t$. How can you tell from this time series of values whether $S$ is chaotic? In this section we give four qualitative criteria:

- the time series "looks chaotic"
- the Poincare map is space-filling
- the power spectrum exhibits broadband noise
- the autocorrelation function decays rapidly

Time Dependence

Using the time series method is both intuitive and easy. The gross behavior of a system can often be learned merely by studying the temporal behavior of each of its variables. The system is likely to be chaotic if such temporal plots are nonrecurrent and appear jagged and irregular. Moreover, sensitivity to initial conditions can be easily tested by simultaneously plotting two trajectories of the same system but starting from slightly different initial states. Figure 8, for example, shows the divergence of two trajectories for the logistic map with $\alpha = 4$.
whose initial points \( x_0 = 0.12345 \) and \( x_0' = 0.12346 \) differ only in the 5th decimal place.

Figure 8. Divergence of trajectories for two nearby initial points (differing by \( x_0 - x_0' = 10^{-4} \)) for the logistic equation for \( \alpha = 4.0 \).

**Poincare Maps**

Recall that the Poincare map is a method for visualizing continuous trajectories by constructing an equivalent discrete-time mapping by periodic "stroboscopic" sampling of points along an orbit. Consider a two-dimensional trajectory in three-dimensional space. The structure of such a trajectory can be readily identified by plotting its intersections with a two-dimensional slice through the three-dimensional space in which it lives. The system is likely to be chaotic if the discrete point set on the resulting Poincare plot is fractal or space-filling.

**Autocorrelation Function**

The autocorrelation function, \( C(\tau) \), of a time series measures the degree to which one part of the trajectory is correlated with itself at another part. If a series is completely random in time, then different parts of the trajectory are completely uncorrelated and the autocorrelation function approaches zero. Put another way, no part of the trajectory harbors any useful information for predicting any later part of the trajectory. As the correlation between parts of a trajectory increases, parts of a trajectory contain an increasing amount of
information that can be used to predict later parts and the value of the autocorrelation function thus increases.

For continuous signals, the autocorrelation function \( C(\tau) \) is defined by

\[
C(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{0}^{T} c(t)c(t + \tau) \, dt,
\]

where \( c(t) = \zeta(t) - \lim_{T \to \infty} \frac{1}{T} \int_{0}^{T} \zeta(t) \, dt. \)

For discrete systems, \( C(\tau) \) is defined by

\[
C(\tau) = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} c(t_i)c(t_i + \tau),
\]

where \( c(t_i) = \zeta(t_i) - \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} \zeta(t_j). \)

A rapid decay (say, with an exponential dropoff) of \( C(\tau) \) is a criterion for the presence of chaos.

**Power Spectrum**

If a system is chaotic that means its signal is irregular and aperiodic in time. A measure that is often used to distinguish between multiply periodic behavior (that can also appear irregular and complicated) and chaos is the power spectrum of the signal, \( P(\omega) \). For continuous systems it is defined by

\[
P(\omega) = |\zeta(\omega)|^2,
\]

where \( \zeta(\omega) \) is the Fourier-transform of the signal \( \zeta(t) \):

\[
\zeta(\omega) = \lim_{T \to \infty} \int_{0}^{T} \zeta(t) e^{i\omega t} \, dt.
\]

For discrete systems, the Fourier-transform of the signal \( \zeta(t) \) is defined by

\[
\zeta_k = \frac{1}{N} \sum_{n=0}^{N-1} \zeta(t) e^{2\pi i nk/N}.
\]

In either case, for multiply periodic motion the power spectrum consists only of discrete lines of the corresponding frequencies. Chaotic motion, on the other hand, induces broadband noise in \( P(\omega) \), that is mostly concentrated in the lower frequencies.
Quantitative Characterization of Chaos

The previous section introduced several qualitative criteria for the presence of chaos in a dynamical system. In this section we define a set of quantitative measures of chaos:

- Lyapunov exponents
- generalized fractal dimensions
- Kolmogorov-Sinai entropy

**Lyapunov Exponents**

As has been repeatedly stressed, a fundamental property of chaotic motion is sensitivity to small changes to initial conditions. Initially closely separated starting conditions evolving along regular dynamical trajectories diverge only linearly in time. A chaotic evolution, on the other hand, leads to an exponential divergence in time. Lyapunov exponents quantify this divergence by measuring the mean rate of exponential divergence of initially neighboring trajectories.

Consider two initial points of a one-dimensional trajectory – \( x_0 \) and \( x'_0 = x_0 + \varepsilon \) – separated by some small quantity \( \varepsilon \). Suppose each of these points evolves according to the map \( x_{n+1} = f(x_n) \), for some function \( f \). Figure 9 shows that after \( N \) steps, the Lyapunov exponent \( \lambda(x_0) \) measures the exponential separation between the \( N \)th iterates of \( x_0 \) and \( x'_0 \), or between \( f^N(x_0) \) and \( f^N(x_0 + \varepsilon) \), respectively.

Figure 9. Schematic definition of Lyapunov exponent

From figure 9, we see that

\[ \varepsilon e^{N\lambda(x_0)} = |f^N(x_0 + \varepsilon) - f^N(x_0)| \]
which, in the limit as \( \varepsilon \to 0 \) and \( N \to \infty \) yields the following expression for \( \lambda(x_0) \):

\[
\lambda(x_0) = \lim_{N \to \infty, \varepsilon \to 0} \frac{1}{N} \log \left| \frac{f'(x_0 + \varepsilon) - f'(x_0)}{\varepsilon} \right|
\]

\[
= \lim_{N \to \infty} \frac{1}{N} \log \left| \prod_{i=0}^{N-1} f'(x_i) \right| = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} \log |f'(x_i)|,
\]

where \( f'(x_0) \) is the derivative of the function \( f \) evaluated at the point \( x_0 \). Thus \( e^{\lambda(x_0)} \) is the average factor by which the distance between initially closely separated points becomes stretched after one iteration. If \( \lambda \leq 0 \), nearby trajectories tend to converge rather than diverge and the motion is regular; if \( \lambda > 0 \), nearby trajectories tend to diverge from one another and the motion is chaotic.

Figure 10. Lyapunov exponent vs. control parameter \( \alpha \) for the logistic equation

As an example, consider the logistic map, defined by \( x_{n+1} = \alpha x_n (1 - x_n) \). A straightforward calculation shows that \( \lambda = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} \log |1 - 2x_i| \). Figure 10 shows a plot of \( \lambda \) vs. \( \alpha \) for \( \alpha \) in the range 2.9 - 4.0. Consistent with our earlier observations of the behavior of this map (see figure 6), we see that (1) \( \lambda \leq 0 \) for all \( \alpha < \alpha_{\infty} \), (2) \( \lambda > 0 \) for most \( \alpha > \alpha_{\infty} \), and (3) that there are multiple windows in the chaotic regime for which \( \lambda \) dips down below zero and the attractor thus becomes periodic.
An n-dimensional system has n one-dimensional Lyapunov exponents, \( \lambda_1, \lambda_2, ..., \lambda_n \). Each \( \lambda_i \) measures the rate of divergence in the \( i^{th} \) direction.

**Information Theoretic Interpretation**

As defined above, the Lyapunov exponents effectively determine the degree of chaos that exists in a dynamical system by measuring the rate of the exponential divergence of initially closely neighboring trajectories. A suggestive alternative interpretation is an information-theoretic one. It is, in fact, not hard to see that Lyapunov exponents are very closely related to the rate of information loss in a dynamical system.

Consider, for example, a one-dimensional interval \([0,1]\), that is partitioned into \( N \) equal sized bins. Assuming that a point \( x_0 \) is equally likely to fall into any one of these bins, learning which bin in fact contains \( x_0 \) therefore constitutes an information gain

\[
I_0 = -\sum_{i=1}^{N} \frac{1}{N} \log_{2} \frac{1}{N} = \log_{2} N 
\]

where \( \log_2 \) is the logarithm to the base 2. Now consider a simple linear one-dimensional map \( f(x) = \alpha x \), where \( x \) is in the interval \([0,1]\) and \( \alpha > 1 \). By changing the length of the interval, and thereby decreasing the effective resolution, by a factor \( \alpha = |f'(0)| \), a single application of the map \( f(x) \) results in an effective information loss

\[
\delta I = I_1 - I_0 = -\sum_{i=1}^{N} \frac{\alpha}{N} \log_{2} \frac{\alpha}{N} + \sum_{i=1}^{N} \frac{1}{N} \log_{2} \frac{1}{N} = -\log_{2} |f'(0)|. 
\]

Generalizing to the case when \( |f'(x)| \) depends on position, and averaging over a large number of iterations, we obtain the following expression for the mean information loss:

\[
\delta I_{ave} = -\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N-1} \log_{2} |f'(x_i)| = \lambda \log_{10} 2, 
\]

where \( \lambda \) is the Lyapunov exponent. We thus see that, in one dimension, \( \lambda \) measures the average loss of information about the position of a point in \([0,1]\) after one iteration.

**Numerical Computation**

There are several useful methods for computing the Lyapunov exponents from experimental time series data, including the so-called "pullback-technique" by Benettin, et. al. [22], a method proposed by Eckman and Ruelle [85] and an algorithm that is particularly well
suited for the analysis of experimental data suggested by Wolf and Vastano [320]. To illustrate the general methodology we briefly discuss the latter algorithm.

The first step is to construct an attractor from the experimental data using the embedding technique. That is, construct from a time series \( \{ \xi(t) \} \) a set of points of the form

\[
\tilde{x}_i = (\xi(t_i), \xi(t_i - \tau), ..., \xi(t_i - (m-1)\tau)),
\]

where \( \tau \) is a fixed time delay. This time-delayed embedding reconstruction assumes that if the embedding dimension \( m \) is large enough, the behavior of whatever system is responsible for generating the particular series of measurements can be described by a finite dimensional attractor. In principle, the choice of \( \tau \) is arbitrary, though criteria for its selection exist.

Figure 11 shows a schematic illustration of the Wolf algorithm to compute the largest Lyapunov exponent, \( \lambda_{max} \).

The Wolf algorithm follows a pair of initially nearby points on the attractor. Begin with a data point \( y(t_0) \) and its nearest neighbor \( z_0(t_0) \), which are a distance \( d_0 = |z_0(t_0) - y(t_0)| \) apart. These two points are evolved by time increments \( \Delta t \) until the distance \( d_0' \) between them exceeds some threshold value \( \epsilon \). When that occurs, the first incremental data point \( y(t_1) \) is retained and a new neighbor \( z_1(t_1) \) is sought such that the distance \( d_1 = |y(t_1) - z_1(t_1)| \) is again less than \( \epsilon \) and such that \( z_1(t_1) \) lies as closely as possible in the same direction from \( y(t_1) \) as \( z_0(t_0) \).

Figure 11. Schematic illustration of the Wolf algorithm for computing the largest Lyapunov exponent
This procedure is continued until the fiducial trajectory \(\overline{y}\) has been followed to the end of the time series. The largest Lyapunov exponent of the attractor, \(\lambda_{\text{max}}\), is then estimated as

\[
\lambda_{\text{max}} = \frac{1}{N_M} \sum_{i=0}^{M-1} \log_2 \frac{d_i}{d_i^*},
\]

where \(M\) is the number of replacements and \(N\) is the total number of time steps for which the fiducial trajectory \(\overline{y}\) has been followed. The presence of chaos in a time series can now be confirmed by finding that \(\lambda_{\text{max}} > 0\). In practice, a few thousand attractor points suffice to estimate \(\lambda_{\text{max}}\) to within 10% of the true value when the attractor is less than three dimensional [320].

**Fractal Dimensions**

While Lyapunov exponents, as discussed in the last section, confirm the presence of chaos by quantifying the magnitude of the exponential divergence of initially neighboring trajectories, they do not provide any useful structural or statistical information about a strange attractor. Such information is instead provided by various fractal dimensions.

Recall that fractals are geometric objects characterized by some form of self-similarity; that is, parts of a fractal, when magnified to an appropriate scale, appear similar to the whole. Fractals are thus objects that harbor an effectively infinite amount of detail on all levels. Coastlines of islands and continents and terrain features are approximate fractals. A magnified image of a part of a leaf is similar to an image of the entire leaf. Strange attractors also typically have a fractal structure.

Loosely speaking, a fractal dimension specifies the minimum number of variables that are needed to specify the fractal. For a one-dimensional line, for example, say the x-axis, one piece of information, the x-variable, is needed to specify any position on the line. The fractal dimension of the x-axis is said to be equal to 1. Similarly, two coordinates are needed to specify a position on a two-dimensional plane, so that the fractal dimension of a plane is equal to 2. Genuine (i.e. interesting) fractals are objects whose fractal dimension is noninteger.

**Box Dimension**

Consider the simplest example of a fractal dimension, sometimes called the box dimension, \(D_{\text{box}}\). Suppose we want to compute the box dimension for a set of points in a d-dimensional space. Define \(N(\varepsilon)\) to be the minimum number of d-dimensional cubes of volume \(\varepsilon^d\) that are
necessary to completely cover the set. The box dimension is then defined as

\[ D_{\text{box}} = \lim_{\varepsilon \to 0} \frac{\log[N(\varepsilon)]}{\log(1/\varepsilon)}. \]

\( D_{\text{box}} \) essentially tells us how much information is needed to specify the set to within an accuracy \( \varepsilon \). In practice, one obtains values of \( N(\varepsilon) \) for a variety of \( \varepsilon \)'s and estimates \( D_{\text{box}} \) from the slope of a plot of \( \log[N(\varepsilon)] \) versus \( \log(1/\varepsilon) \).

This expression for \( D_{\text{box}} \) gives the expected result for simple sets. If the set in question consists of a single point, for example, we know that one box of any size \( \varepsilon > 0 \) suffices to cover the set, so that \( D_{\text{box}} = \log(1)/\log(1/\varepsilon) = 0 \), as expected. Similarly, if the set in question is a line segment of length \( L=1 \), then we can take \( N(\varepsilon) = 1/\varepsilon \) so that \( D_{\text{box}} = \log(1/\varepsilon)/\log(1/\varepsilon) = 1 \). In fact, for the usual d-dimensional Euclidean sets, the box dimension equals the topological dimension. One of the simplest examples of a set for which the two measures differ is the so-called Cantor fractal.

Figure 12 shows the first four steps in the construction. The first step consists of a line of length \( L=1 \). Call this set \( S_1 \). At the next step, set \( S_2 \) is obtained by deleting from \( S_1 \) its middle third. At the third step, the set \( S_3 \) is obtained from \( S_2 \) by deleting the middle third segments from each of the two disjoint pieces making up \( S_2 \). Continue in this fashion, at each step \( n \) deleting the middle third segments from each of the disjoint pieces making up the set obtained on the previous step, \( n-1 \). The Cantor fractal is the set that remains in the limit as \( n \to \infty \).

**Figure 12. First four steps of the construction of the Cantor fractal**

- **step 1:**
- **step 2:**
- **step 3:**
- **step 4:**

Despite the fact that the length of the Cantor fractal is zero –

\[ L = 1 - \frac{1}{3} - \frac{2}{9} - \frac{4}{27} - \cdots = 1 - \frac{1}{3} \sum_{i=0}^{\infty} \left( \frac{2}{3} \right)^i = 0 \]
- its box dimension is greater than zero. Since at the $n^{th}$ step of the construction $N(\varepsilon) = 2^n$ balls of size $\varepsilon = (1/3)^n$ are needed to cover the set, we see that

$$D_{box}(\text{Cantor}) = \lim_{n \to \infty} \frac{\log(2^n)}{\log(3^n)} \approx 0.6309.$$  

As another example, consider the Feigenbaum attractor of the logistic equation (see page 34) for $\alpha = \alpha_\infty \approx 3.56994\ldots$ One can show that the trajectory of points at this value of $\alpha$ is a fractal with box dimension $D_{box} \approx 0.5388$.

Note that while $D_{box}$ clearly depends on the metric properties of the space in which the attractor is embedded – and thus provides some structural information about the attractor – it does not take into account any structural inhomogeneities in the attractor. In particular, since the box bookkeeping only keeps track of whether or not an overlap exists between a given box and the attractor, the individual frequencies with which each box is visited are ignored. The inhomogeneities of the Henon attractor, for example (see page 41), and the information that such inhomogeneities might convey about the attractor, are completely ignored by $D_{box}$. This oversight is corrected for by the so-called information dimension, which depends on the visitation frequencies of points on the attractor.

**Information Dimension**  

Just as for the box dimension, first partition the $d$-dimensional space into boxes of volume $\varepsilon^d$. The probability of finding a point of an attractor in box number $i$, where $i = 1, 2, \ldots, N(\varepsilon)$, is $p_i(\varepsilon) = N_i(\varepsilon)/N$, where $N_i(\varepsilon)$ is the number of points in the $i$th box and $N$ is the total number of points. $p_i(\varepsilon)$ is thus the relative frequency with which the $i^{th}$ box is visited, and ranges from zero (when $N_i(\varepsilon) = 0$) to one (when $N_i(\varepsilon) = N$).

The amount of information required to specify the state of the system to within an accuracy $\varepsilon$ (or, equivalently, the information gain in making a measurement that is uncertain by an amount $\varepsilon$) is given by

$$I(\varepsilon) = -\sum_{i=1}^{N(\varepsilon)} p_i(\varepsilon) \log p_i(\varepsilon).$$

The information dimension, $D_I$, is then defined as

$$D_I = \lim_{\varepsilon \to 0} \frac{I(\varepsilon)}{\log(1/\varepsilon)}.$$  

Notice that if the set is contained entirely in a single box – say, the 13$^{th}$ box – then $p_{13}(\varepsilon) = 1$, and $p_i(\varepsilon) = 0$ for all $i \neq 13$. Thus $D_I = 0$. On the
other hand, if each box is visited equally often – that is, if $p_i(\varepsilon) = 1/N(\varepsilon)$ for all $i$ – than $I(\varepsilon) = \log [N(\varepsilon)]$ and $D_I = D_{\text{box}}$. For unequal probabilities, $I(\varepsilon) < \log [N(\varepsilon)]$, so that, in general, $D_I \leq D_{\text{box}}$.

**Correlation Dimension**

Another important measure is based on the correlation integral $C(\varepsilon)$ introduced by Grassberger and Procaccia in 1983. $C(\varepsilon)$ measures the probability of finding two points of an attractor in a box of size $\varepsilon$:

$$
\sum_{i=0}^{B(\varepsilon)} p_i^2 = \text{the probability that two points lie within the box } \varepsilon^d
$$

$$
\equiv \text{the probability that two points are separated by a distance smaller than } \varepsilon
$$

$$
= \lim_{N \to \infty} \frac{1}{N^2} \sum_{i,j} \theta (\varepsilon - |\vec{x}_i - \vec{x}_j|)
$$

$$
= C(\varepsilon) = \text{correlation integral},
$$

where $\theta(x)$ denotes the Euclidean distance, and $\theta(x) = 1$ for $x > 0$ and $\theta(x) = 0$ otherwise. $C(\varepsilon)$ essentially counts the number of pairs of points falling within a hypersphere of radius $\varepsilon$ that is centered on each point (and normalizes by a factor $1/N^2$). The **correlation dimension**, $D_{\text{corr}}$, is then defined as

$$
D_{\text{corr}} = \lim_{\varepsilon \to 0} \frac{\log [C(\varepsilon)]}{\log (1/\varepsilon)} = \lim_{\varepsilon \to 0} \frac{\log \left( \sum p_i^2 \right)}{\log (1/\varepsilon)}.
$$

It can be shown that $0 \leq D_{\text{corr}} \leq D_I \leq D_{\text{box}}$.

The correlation integral and correlation dimension can be used to determine two additional properties from experimental time series data:

- **the embedding dimension $d_e$**: the dimension $d_e$ in the time series $\vec{x}(t) = \{ \zeta(t_1), \zeta(t_1 + \tau), ..., \zeta(t_1 + (d_e - 1)\tau) \}$ above which $D_{\text{corr}}$ no longer changes is the (minimal) embedding dimension of the attractor.

- **distinction between deterministic chaos and random noise**: suppose there is a strange attractor embedded in $d$-dimensional space and an external random noise is added. Each point on the attractor becomes surrounded by a uniform $d$-dimensional
cluster of points. Suppose the radius of this cluster is $\epsilon_{\text{noise}}$. Then for $\epsilon >> \epsilon_{\text{noise}}$, the correlation integral $C(\epsilon)$ counts these clusters as points and the slope of the log $\log \left( [C(\epsilon)] \right)$ versus $\log(\epsilon)$ plot yields the correlation dimension of the attractor. For $\epsilon < \epsilon_{\text{noise}}$, most of the points counted by $C(\epsilon)$ fall within the uniformly filled d-dimensional clusters and the slope of the log $\log \left( [C(\epsilon)] \right)$ versus $\log(\epsilon)$ plot crosses over to d.

**Hierarchy of Generalized Fractal Dimensions**

The three fractal dimensions discussed in the previous section – the box dimension, $D_{\text{box}}$, the information dimension $D_I$, and the correlation dimension, $D_{\text{corr}}$ – are actually three members of an infinite hierarchy of generalized fractal dimensions, $D_q$, that characterize an attractor:

$$D_q = \lim_{\epsilon \to 0} \frac{B(\epsilon)}{\log \left( \sum_{i=1}^{\infty} p_i^q(\epsilon) \right)} \frac{1}{1-q} \frac{\log(\epsilon)}{\log(1/\epsilon)},$$

where $p_i(\epsilon)$ is, as before, the probability of finding a point of an attractor in the $i$th box of size $\epsilon$, where $i = 1, 2, ..., B(\epsilon)$.

It can be shown that $D_{q=0}$ corresponds to the box dimension, $D_{q=1}$ to the information dimension, and $D_{q=2}$ to the correlation dimension. It can also be shown that $D_q$ is a non-decreasing function of $q$: i.e. that $D_q \leq D_{q'}$ for all $q$, $q'$ such that $q \geq q'$.

**Lyapunov Dimension**

An attempt to link a purely static property of an attractor – as embodied by its box dimension, $D_{\text{box}}$ – to a dynamic property, as expressed by its set of Lyapunov exponents, $\{\lambda_i\}$, was first made by Kaplan and Yorke [168]. Defining the Lyapunov dimension, $D_L$, to be

$$D_L = j + \frac{\sum_{i=1}^{j} \lambda_i}{|\lambda_{j+1}|},$$

where $j$ denotes the largest $l$ such that $\sum_{i=1}^{j} \lambda_i \geq 0$, the Kaplan-Yorke conjecture is that $D_{\text{box}} = D_L$. Although this equality seems to be satisfied exactly only for completely homogeneous attractors, it is often approximately satisfied by inhomogeneous attractors as well. Because the calculation of the Lyapunov exponents $\lambda_i$ is relatively easy, this
simple relation has proven to be useful for obtaining quick characterizations of strange attractors.

**Kolmogorov-Sinai Entropy**

The discussion has so far focused exclusively on the amount of information gained from a single 'snap-shot' view of an attractor. A useful alternative viewpoint is provided by asking about the rate of information gain per unit time achieved in observing the system over a period of time. If the dynamics are simple, the asymptotic information gain is zero since new measurements provide no new information; if the system is behaving chaotically, on the other hand, new measurements are constantly needed in order to update our knowledge of the system. The Kolmogorov-Sinai entropy (or metric entropy), $K$, gives an upper bound on this information acquisition rate.

To define $K$, partition a $d$-dimensional phase space into boxes of volume $\varepsilon^d$, $\{b_1, b_2, ..., b_N\}$. Assuming that the state of the system is measured in intervals $\Delta t = \tau$ units, define $p(b_1, b_2, ..., b_N)$ to be the joint probability that $\tilde{x}(t = 0)$ is in box $b_1$, $\tilde{x}(t = \tau)$ is in box $b_2$, ..., and $\tilde{x}(t = N\tau)$ is in box $b_N$. The Shannon information, $I_N$, stored in this string is proportional to the information needed to locate the system on a trajectory $(b_1, b_2, ..., b_N)$ with precision $\varepsilon$ (if one knows a priori only the probabilities $p(b_1, b_2, ..., b_N)$), and is given by

$$I_N = - \sum_{b_1, b_2, ..., b_N} p(b_1, b_2, ..., b_N) \log p(b_1, b_2, ..., b_N).$$

The additional information needed to predict which box $\tilde{x}(t = (N+1)\tau)$ will be in, given $I_N$, is given by $I_{N+1} - I_N$ which is therefore a measure of our information gain about the state of the system from time $t = N\tau$ to $t = (N+1)\tau$. The Kolmogorov-Sinai entropy is the average rate of this information gain:

$$K = \lim_{\tau \to 0} \lim_{\varepsilon \to 0} \lim_{M \to \infty} \frac{1}{M\tau} \sum_{N=0}^{M-1} (I_{N+1} - I_N),$$

$$K = \lim_{\tau \to 0} \lim_{\varepsilon \to 0} \lim_{M \to \infty} \frac{1}{M\tau} \sum_{b_1, b_2, ..., b_{M-1}} p(b_1, b_2, ..., b_{M-1}) \log p(b_1, b_2, ..., b_{M-1}).$$

The limit $\varepsilon \to 0$ (which must be taken after the $M \to \infty$ limit) makes $K$ independent of a particular partitioning of the phase space. For discrete maps with discrete time steps $\tau = 1$, the limit $\tau \to 0$ is omitted. It is easy to see that $K = 0$ for regular trajectories, while completely random motion yields $K = \infty$. Deterministic chaotic motion, on the other hand, results in $K$ being both finite and positive. A method for
deriving $K$ of strange attractors from scalar time series is given by Fraser [100].

**Time-Series Forecasting and Predictability**

As has been repeatedly stressed throughout this discussion, chaos theory tells us that a chaotic dynamical system is sensitive to initial conditions. This, in turn, implies that chaos precludes long-term predictability of the behavior of the system. The essence of chaos, after all, is the unpredictability of individual orbits; thinks of the random sequence of heads and tails from tosses of an unbiased coin or the dripping of a faucet. On the other hand, suppose a system's orbit lies on a strange attractor. If we know something about this attractor – its general shape, for example, perhaps along with an estimate of the visitation frequencies to its different parts – this clearly provides some information about what the deterministic (albeit chaotic) system is doing. This added information, in turn, may be sufficient to allow us to make predictions about certain short-term (and long-term) behavioral trends of the system.

Chaotic dynamics is often *misinterpreted* to mean *random* dynamics. Strictly speaking, since chaos is spawned from a deterministic process, its apparent irregularity stems from an intrinsic magnification of an external uncertainty, such as that due to a measurement of initial conditions. Sensitivity to initial conditions amplifies an initially small uncertainty into an exponentially large one; or, in other words, short-term determinism evolves into long-term randomness. Thus, as Eubank and Farmer\(^6\) point out, the important distinction is not between chaos and randomness, but between chaotic dynamical systems that have low-dimensional attractors and those that have high-dimensional attractors. For example, if a time series of evolving states of a system is generated by a very high dimensional attractor (or if the dynamics is modeled in a state space whose dimension is less than that of the attractor), then it will be essentially impossible to gather enough information from the time series to exploit the underlying determinism. In this case, the apparent randomness will in fact have become a very real randomness, at least from a predictability standpoint. On the other hand, if the time series is generated by a relatively low dimensional attractor, it is possible to exploit the underlying determinism to predict certain aspects of the overall behavior. A powerful technique to make the underlying determinism of a chaotic time series stand out is the so-called *embedding technique.*

\(^6\) S. Eubank and D. Farmer, "An introduction to chaos and randomness," pages 75-190 in [159].
State-Space Reconstruction via Embedding

Consider some real-world data, tabulated as a time series,
\[ \tilde{\zeta} = \{\xi(t_1), \xi(t_2), \ldots, \xi(t_N)\} \]. The data may represent observations of the closing prices of the Dow Jones Industrials, annual defense expenditures, or combat losses on the battlefield.

The embedding technique is a method of reconstructing a state space from the time series. It assumes that if the embedding dimension is large enough, the behavior of whatever system is responsible for generating the particular series of measurements can be described by a finite dimensional attractor. Its main strength lies in providing detailed information about the behavior of degrees-of-freedom of a system other than the ones that are directly observed. Estimates of the error introduced by extrapolating the data can also be made.

The embedding technique consists of creating the state vectors \( \tilde{x}_i \) from \( \tilde{\zeta} \) according to

\[ \tilde{x}_i = (\xi(t_i), \xi(t_i + \tau), \ldots, \xi(t_i + (m-1)\tau)) \]

where \( \tau \) is a fixed time delay. In principle, the choice of \( \tau \) is arbitrary, though criteria for its selection exist. If the dynamics takes place on an attractor of dimension \( d \), then a necessary condition for "uncovering" the underlying determinism is \( m \geq d \). It can be shown that if \( r \) is the dimension of a manifold containing the attractor, than almost any embedding in \( d = 2r + 1 \) dimensions will preserve the topological properties of the attractor. Of course, the embedding technique does not work for all time series, and the amount of information it uncovers about the underlying determinism for a given time series may be sufficient only to yield very short-term predictions. Nonetheless, the technique has proven to be very powerful in uncovering patterns in data that are not otherwise (obviously) visible. A detailed discussion is given in reference [54].

Figure 13 shows an example of the kind of predictions that are possible with the embedding technique. Given 1000 data points (not shown) of the chaotic fluctuations in a far-infrared laser (approximately described by three coupled nonlinear ordinary differential equations) from which to learn the underlying system's dynamics, Sauer [275] uses a modified embedding technique to predict the continuation of the time series for 200 additional time steps. Figure 13 (a-d) shows four continuations of length 200, each with a different initial point. In each of the plots, the solid curve represents the predicted continuation, and the dashed curve represents the true continuation.
Figure 13. Four continuations of a chaotic time series using the embedding technique; solid lines represent predicted values, dashed lines represent the actual data.

Chaotic Control

Suppose you have a physical system that exhibits chaos. Is there a way to still use the system — that is, to allow the system to evolve naturally according to its prescribed dynamics — but in such a way as to eliminate that system's chaotic behavior? One way, of course, might be to physically alter the system in some (possibly costly) way. But what if such a restructuring is not an option? What if the only available option is to slightly "tweak" one of the system's control parameters?

It has recently been shown by Ott, et. al. [237] and Romeiras, et. al. [268] that the extreme sensitivity of chaotic systems to small perturbations to initial conditions (the so-called "butterfly effect") can be exploited to stabilize regular dynamic behaviors and to effectively "direct" chaotic trajectories to a desired state. The critical idea is that
chaotic attractors typically have embedded within them a dense set of unstable periodic orbits. That is to say, an infinite number of unstable periodic orbits typically co-exist with the chaotic motion. By a periodic orbit, we mean an orbit that repeats itself after some finite time. If the system were precisely on an unstable periodic orbit, it would remain there forever. Such orbits are unstable because the smallest perturbation from the periodic orbit (as might, for example, be due to external random noise) is magnified exponentially in time and the system orbit moves rapidly away from the periodic orbit. The result is that while these unstable periodic orbits are always present, they are not usually seen in practice. Instead, one sees a chaotic orbit that bounces around in an apparently random fashion. Ironically, chaotic control is a capability that has no counterpart in nonchaotic systems. The reason is that the trajectories in nonchaotic systems are stable and thus relatively impervious to desired control.

The basic strategy consists of three steps:

- find some unstable periodic orbits embedded within the chaotic motion
- examine these orbits to find an orbit that yields an improved system performance
- apply small controlling perturbations to direct the orbit to the desired periodic (or steady state) motion

Once a desired unstable periodic orbit has been selected, the nature of a chaotic orbit assures us that eventually the random-appearing wanderings of the chaotic orbit will bring it close to the selected unstable periodic orbit. When this happens the controlling perturbations can be applied. Moreover, if there is any noise present, these controlling perturbations can be applied repeatedly to keep the trajectory on the desired orbit.

We make a few general comments:

1. Chaotic control is applicable to both continuous and discrete dynamical systems.

2. Chaos can be controlled using information from previously observed system behavior. Thus it can be applied to experimental (i.e. real-world) situations in which no model need be available to define the underlying dynamics.

3. While chaotic control applies strictly to systems that are described with a relatively few variables, it should be
remembered that the behavior of many high (and even infinite) dimensional systems is often described by a low dimensional attractor.

4. Before settling into a desired controlled orbit the trajectory goes through a chaotic transient whose duration diverges as the maximum allowed size of the control perturbations approaches zero.

5. Small noise can result in occasional bursts in which the orbit strays far from the desired orbit.

6. Any number of different orbits can be stabilized, where the switching from one to another orbit is regulated by corresponding control perturbations.

A recent survey article [284] lists applications for communications (in which chaotic fluctuations can be put to use to send controlled, pre-planned signals), for physiology (in which chaos is controlled in heart rhythms), for fluid mechanics (in which chaotic convection currents can be controlled) and chemical reactions. As another recent example, a few years ago NASA used small amounts of residual hydrazine fuel to steer the ISEE-3/ICE spacecraft to its rendezvous with a comet 50 million miles away. This was possible because of the sensitivity of the three-body problem of celestial mechanics to small perturbations.

**Brief Overview of Method**

Consider a discrete d-dimensional dynamical system, \( \hat{Z}_{n+1} = \hat{F}(\hat{Z}_n, p) \), where \( \hat{Z}_n \) is a d-dimensional vector describing the state of the system at time-step "n" and \( p \) is a control parameter (which will be used for inducing "control perturbations"). This control parameter is adjustable but is restricted to within a range \( p^\prime - \delta < p < p^\prime + \delta \), where \( p^\prime \) is the nominal value for which the system has a chaotic attractor, and \( \delta \) is some small number. The problem is now to vary \( p \) in such a way that for almost all initial conditions in the basin of the chaotic attractor, the system will converge onto the desired periodic orbit embedded within the attractor. For simplicity, we focus attention on stabilizing fixed point (i.e. period one) orbits; generalization to higher period orbits is straightforward [268].

The first step is to approximate the dynamics near the fixed point, labeled \( \hat{Z}^* \) (so that \( \hat{Z}^* = \hat{F}(\hat{Z}^*, p^*) \)). For values of \( p \) close to \( p^\prime \), this approximation is given by the linear map.
\[
(\dot{Z}_{n+1} - \dot{Z}^*) = A \cdot (\dot{Z}_n - \dot{Z}^*) + B(p - p^*),
\]

where \( A = \frac{\partial \tilde{F}}{\partial \tilde{Z}} \) is a d-by-d dimensional jacobian matrix and \( B = \frac{\partial \tilde{F}}{\partial p} \) is a d-dimensional column vector. Both \( A \) and \( B \) are evaluated at \( Z = Z^* \) and \( p = p^* \).

Assuming that we can tweak the parameter \( p \) on each iteration, we replace \( p \) by \( p_n \) according to the following linear prescription:

\[
p \rightarrow p_n = p^* - K^T \cdot (\dot{Z}_n - \dot{Z}^*),
\]

where \( K \) is a constant d-dimensional column vector, and \( K^T \) is its transpose. The 1-by-d matrix \( K^T \) must be determined so that the fixed point \( Z^*(p^*) \) becomes stable. Substituting this value of \( p_n \) into the above expression for \( (\dot{Z}_{n+1} - \dot{Z}^*) \) we find

\[
\delta Z_{n+1} = (A - B \cdot K^T) \delta Z_n,
\]

where \( \delta \dot{Z}_n = \dot{Z}_n - \dot{Z}^* \). It is clear that the fixed point will be stable if the matrix \( A - B \cdot K^T \) only has eigenvalues whose modulus is less than one. The "pole placement solution" to this problem is well known in the literature, and is summarized by Romeiras, et. al. [268].

**Lessons of Nonlinear Dynamics and Chaos**

The major lesson of nonlinear dynamics is that a dynamical system does not have to be "complex" or to be described a large set of equations, in order for the system to exhibit chaos – all that is needed is three or more variables and some embedded nonlinearity.

Basic lessons of nonlinear dynamics and deterministic chaos include:

- chaos is pervasive – apparently random behavior in some nonlinear systems can in fact be described by deterministic (non-random) chaos
- nonlinear dynamics teaches us to appreciate the wide range of qualitatively different dynamical behaviors that can be generated by feedback in real systems
- nonlinear systems generally tend to exhibit bifurcations – small changes in parameters can lead to qualitative transitions to new types of behaviors
• small perturbations can induce large changes

• typical nonlinear systems have multiple basins of attraction, and the boundaries between different basins can have very complicated fractal forms

• dynamical behavior depends on location in phase space

• an appreciation of what transitions to expect when one adds feedback to a system

• suggest ways in which to selectively adjust feedback

• an understanding that while individual trajectories behave in an apparently erratic manner, the attractors themselves offer information about the long-term trends of a system

• techniques such as time-delayed embedding allow short-term prediction even without any prior knowledge of an underlying model or set of equations

• attractors embody information about certain recurrent aspects of the long-term behavior of a system

• the relative time that an orbit spends visiting various parts of an attractor yields useful visitation probabilities

• the presence of multiple attractors may be exploited for strategic purposes

• the information dimension can be used to estimate the minimal number of variables that are needed to describe the system

• that there are dense paths of trajectories on a chaotic attractor implies that chaos can be exploited to control dynamics that are otherwise erratic and unpredictable

• chaos often results when a dynamical system is not allowed to relax between events

• the universality of certain nonlinear phenomena implies that we may be able to understand many disparate systems in terms of a few simple paradigms and models

Mayer-Kress [38] points out that a failure to learn the lessons of chaos theory could lead to:
• the illusory belief that successful short-term management allows total control of a system

• difficulty, or even impossibility, of making a diagnosis from available short-term data

• application of inappropriate control mechanisms that can actually produce the opposite of a desired effect

Tools for the Decision Maker?

Nonlinear dynamics makes clear that chaotic dynamics ought not be misinterpreted to mean random dynamics. The most important lesson of deterministic chaos is that dynamical behavior that appears to be chaotic or random often contains an embedded regularity. If this embedded regularity can be uncovered and identified, it can potentially be exploited by the decisions maker:

• Short Term Predictions. Given sufficient data, time series analysis permits one to make short-term predictions about a system’s behavior, even if the system is chaotic. Moreover, these prediction can be made even when the underlying dynamics is not known.

• Long-term Trends. If the attractors of a system are known or can be approximated (say, from available historical time series data), long-term trends can be predicted. Knowledge about visitation frequencies of points on an attractor provides insight into the probabilities of various possible outcomes. Lyapunov exponents quantify the limits of predictability.

• Qualitative Understanding of the Battlefield. The information dimension can be used to estimate the minimum number of variables needed to describe a system. Moreover, if a system can be shown to have a small non-integer dimension, it is probable that the underlying dynamics are due to nonlinearities and are not random [231].
Complex systems

"Time is a river which sweeps me along, but I am the river; it is a tiger which destroys me, but I am the tiger; it is a fire which consumes me, but I am the fire." – Jorge Luis Borges

Introduction

What are complex systems? There are examples of complex systems just about everywhere we look in nature, from the turbulence in fluids to global weather patterns to beautifully intricate galactic structures to the complexity of living organisms. All such systems share at least this one property: they all consist of a large assemblage of interconnected, mutually (and typically nonlinearly) interacting parts. Moreover, their aggregate behavior is emergent. That is to say, the properties of the "whole" are not possessed by, nor are they directly derivable from, any of the "parts" – a water molecule is not a vortex, and a neuron is not conscious. A complex system must therefore be understood not just in terms of the set of components out of which it is constructed, but the topology of the interconnections and interactions among those components.

Gases, fluids, crystals, and lasers are all familiar examples of complex systems from physics. Chemical reactions, in which a large number of molecules conspire to produce new molecules, are also good examples. In biology, there are DNA molecules built up from amino acids, cells built from molecules, and organisms built from cells. On a larger scale, the national and global economies and human culture as a whole are also complex systems, exhibiting their own brand of global cooperative behavior. One of the most far-reaching ideas of this sort is James Lovelock's controversial "Gaia" hypothesis, which asserts that the entire earth – molten core, biological ecosystems, atmospheric weather patterns and all – is essentially one huge, complex organism, delicately balanced on the edge-of-chaos.

Perhaps the quintessential example of a complex system is the human brain, which, consisting of something on the order of $10^{10}$ neurons with $10^3 - 10^4$ connections per neuron, is arguably the most "complex" complex system on this planet. Somehow, the cooperative dynamics of this vast web of "interconnected and mutually interacting parts" manages to produce a coherent and complex enough structure for the brain to be able to investigate its own behavior.

The emerging new sciences of complexity and complex adaptive systems explore the important question of whether, or to what extent, does the behavior of the many seemingly disparate complex systems found in nature – from the very small to the very large – stem from the same fundamental core set of universal principles.
References include monographs (Kauffman [171], Holland [144], Mainzer [205], Weisbuch [316]), popularizations (Lewin [196], Waldrop [311], and Gell-Mann [109]), conference proceedings (Cowan, et. al. [60], Varela [307], Yates [326]) and a series of lecture notes from the Santa Fe Institute ([159], [227]-[230]).

**Short History**

Table 4 shows a brief chronology of some of the milestone events in the study of complex systems.

Whenever a new field emerges, many different individuals contribute to its development. This is of course also true of complex systems theory, yet four persons stand out as originating and shaping much of the field: Alan Turing, John von Neumann, Stephen Wolfram and Chris Langton.

Turing, in 1936, published a landmark proof of what has come to be known as the *Halting Theorem*. Turing's theorem fundamentally limits what one is able to know about the running of a program on a computer by asserting that there is in general no way to know in advance if an arbitrary program will ever stop running. In other words, there is, in general, no quick and dirty short-cut way of predicting an arbitrary program's outcome; this is an example of what is called *computational irreducibility*. About five decades later, Wolfram suggested that computational irreducibility is actually a property not just of computers, but of many real physical systems as well.

Cellular automata were conceived in 1948 by John von Neumann, whose motivation was in finding a reductionist model for biological evolution. His ambitious scheme was to abstract the set of primitive logical interactions necessary for the evolution of the complex forms of organization essential for life. In a seminal work, completed by Burks, von Neumann followed a suggestion by Ulam to use discrete rather than continuous dynamics and constructed a two-dimensional automaton capable of self-reproduction. Although it obeyed a complicated dynamics and had a rather large state space, this was the first discrete parallel computational model formally shown to be a *universal computer* (which implies, in turn, that it is also computationally irreducible). Twenty years later, the mathematician John Conway introduced his well-known Life game, which remains among the simplest known models proven to be computational universal.

Other important historical landmarks include the founding, in 1984, of the *Santa Fe Institute*, which is one of the leading interdisciplinary centers for complex systems theory research; the first conference devoted solely to research in cellular automata (which is a prototypical mathematical model of complex systems), organized by Wolfram and...
Land Warfare and Complexity, Part I: Mathematical Background and Technical Sourcebook

Toffoli at MIT in 1986; and the first artificial life conference, organized by Chris Langton at Los Alamos National Laboratory, in 1987.

Table 4. Some historical developments in the study of complex systems

<table>
<thead>
<tr>
<th>Year</th>
<th>Researcher</th>
<th>Discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>1936</td>
<td>Turing</td>
<td>formalized concept of computability; universal turing machine</td>
</tr>
<tr>
<td>1948</td>
<td>von Neumann</td>
<td>wanted to abstract the logical structure of life; introduced self-reproducing automata as a means towards developing a reductionist biology</td>
</tr>
<tr>
<td>1950</td>
<td>Ulam</td>
<td>proposed need for having more realistic models for the behavior of complex extended systems</td>
</tr>
<tr>
<td>1966</td>
<td>Burks</td>
<td>completed and described von Neumann's work</td>
</tr>
<tr>
<td>1969</td>
<td>Zuse</td>
<td>introduced concept of &quot;computing spaces,&quot; or digital models of mechanics</td>
</tr>
<tr>
<td>1970</td>
<td>Conway</td>
<td>introduced two-dimensional cellular automaton Life rule</td>
</tr>
<tr>
<td>1977</td>
<td>Toffoli</td>
<td>applied cellular automata directly to modeling physical laws</td>
</tr>
<tr>
<td>1983</td>
<td>Wolfram</td>
<td>wrote a landmark review article on properties of cellular automata that effectively legitimized the field as research endeavor for physicists</td>
</tr>
<tr>
<td>1984</td>
<td>—</td>
<td>Santa Fe Institute founded, serving as a pre-eminent center for the interdisciplinary study of complex systems</td>
</tr>
<tr>
<td>1986</td>
<td>Toffoli, Wolfram</td>
<td>first cellular automata conference held at MIT, Boston</td>
</tr>
<tr>
<td>1987</td>
<td>Langton</td>
<td>first artificial life conference held at the Santa Fe Institute</td>
</tr>
</tbody>
</table>

Ants and brains ... and combat forces?

Achilles: Familiar to me? What do you mean? I have never looked at an ant colony on anything but the ant level.

Anteater: Maybe not, but ant colonies are no different from brains in many respects...

— Douglas Hofstadter, Godel, Esher, Bach

Much has been written about how insect "societies" — with their complex hierarchies of function and responsibility — often exhibit intelligent-like behavior. Consider the massive mounds built by the termite Macrotermes. The heat generated within these mound is carried upwards via a central air duct where it then travels back down along narrow channels lying close to the surface and where it is cooled and oxygen and carbon dioxide are exchanged, just as in a lung. Such mounds, as whole, act as air-conditioning system. Although these mounds can be likened to human buildings, in that they are clearly
constructed for the well-being and comfort of its occupants, they are fundamentally different in that they are not engineered. That is to say, no one of their builders ever has any global conception of the structure before it is completed. The mounds emerge from the local-rule governed behavior of tens of thousands of interacting worker termites. The "swarm intelligence" responsible for the structure is itself an emergent collective property of the termite society as a whole, and is a property clearly not possessed by any of the society's non-intelligent constituents.

Just as ant and termite colonies and brains share many nontrivial collective properties, it can also be argued that there are strong analogies between these "social mind" systems and the self-organizing dynamics of combat forces:

- the behavior of individual elements (whether they be ants, neurons or infantrymen) yields little information about the properties or progress of the "collective"

- the global dynamics of each type of system stems from the cooperative nonlinear interaction of individual elements with the environment

- the global behavior of each system is relatively insensitive to the removal of a small number of its individual elements

- each system appears collectively to be "driven," at times, by forces non existent and/or non-acting at their constituent levels (think of the military historian's use of phrases like "shifting-momentum" and "tempo-of-battle" to describe predominantly "high-level" activity on the battlefield)

**Collectivism**

The study of complex systems is not so much a well-developed methodology that comes armed with ready-made IMSL-like algorithms and software routines, as much as a new philosophy, or a new way of looking at some (sometimes very old) problems. The term *collectivism* has sometimes been used to distinguish this philosophy from the more traditional "top down" and "bottom up" philosophies that it embodies.

Collectivism embodies the belief that in order to properly understand complex systems, such systems must be viewed as coherent wholes whose open-ended evolution is continuously fueled by nonlinear feedback between their macroscopic states and microscopic constituents. It is neither completely reductionist (which seeks only to decompose a system into its primitive components), nor completely
synthesist (which seeks to synthesize the system out of its constituent parts but neglects the feedback between emerging levels).

Figure 14. Feedback between local and global levels = Collectivism

As an example of the importance of collectivism, consider a natural ecology. Each species that makes up an ecology composed of a large number of diverse species, co-evolves with other members of the ecology according to a fitness function that is, in part, itself a function of the emerging ecology. Individual members of each species collectively define a (part of the) co-evolving ecology; the ecology, in turn, determines the fitness-function according to which its constituent parts evolve (see figure 14). It is this nonlinear feedback between the information describing individual species (or the system's microscopic level) and the global ecology (or the system's macroscopic level) that those species collectively define that determines the temporal evolution – and identity – of the entire system.

Collectivism is thus distinct from both the top-down reductionist approach traditionally favored by most physicists (system as a simple edifice of its microscopic parts), and the more recent neural-net-like bottom-up approach favored by connectionists (system as a synthesis of its constituent parts). The nonlinear inter-level feedback loop that makes up the collective is what makes a traditional linear analysis of such systems difficult, if not impossible. "Analysis" proceeds from the assumption that in order to understand a system one must first break it up into its constituent parts. Understanding then comes from the knowledge gained by reconstructing the system. But for systems whose dynamics depend critically on interaction between parts, analysis often misses the essential characteristics of the whole system. "Synthesis" is the complementary act of putting the individual pieces together in
order to understand what they do collectively. Understanding complex synthesis requires that both analysis and synthesis be done.

Self-Organization

Self-organization is a fundamental characteristic of complex systems. It refers to the emergence of macroscopic nonequilibrium organized structures, and is due to the collective interactions of the constituents of a complex system as they react and adapt to their environment. There is no God-like "oracle" dictating what each and every part ought to be doing; parts act locally on local information and global order emerges without any need for external control.

At first sight, self-organization appears to violate the Second Law of Thermodynamics, which asserts that the entropy $S$ of an isolated system never decreases (or, more formally, $dS/dt \geq 0$); see figure 15-a. Since entropy is essentially a measure of the degree of disorder in a system, the Second Law is usually interpreted to mean that an isolated system will become increasingly more disordered with time. How, then, can structure emerge after a system has had a chance to evolve?

**Figure 15. Schematic of Isolated and Nonisolated Systems**

![Diagram](attachment:figure15.png)

Upon closer examination, we see that self-organization in complex system does not really violate the Second Law. The reason is that the Second Law requires a system to be isolated; that is, it must not exchange energy or matter with its environment. For nonisolated systems consisting of noninteracting or only weakly interacting particles (see figure 15-b), the entropy $S$ consists of two components: (1) an internal component, $S_i$, due to the processes taking place within the system itself, and (2) an external component, $S_e$, due to the exchange of energy and matter between the system and the environment. The rate of change of $S$ with time, $dS/dt$, now becomes $dS/dt = dS_i/dt + dS_e/dt$. As for an isolated system, $dS_i/dt \geq 0$. But there are no constraints on $dS_e/dt$. If $dS_e/dt$ is sufficiently less than zero, the overall entropy of the system can itself decrease. Thus, the entropy of a nonisolated system of noninteracting or only weakly
interacting particles can decrease due to the exchange of energy and/or matter between the system and its environment.

The situation is more complicated for nonisolated systems consisting of strongly interacting particles and when the system is no longer in equilibrium with the environment. Kauffman [171] notes that the "second law really state that any system will tend to the maximum disorder possible, within the constraints due to the dynamics of the system."

Emergence

Central to the general science of complexity is the concept of emergence. Emergence refers to the appearance of higher-level properties and behaviors of a system that – while obviously originating from the collective dynamics of that system's components – are neither to be found in nor are directly deducible from the lower-level properties of that system. Emergent properties are properties of the "whole" that are not possessed by any of the individual parts making up that whole.

One of the simplest, and ubiquitous, examples of emergence is "temperature," as read by a conventional thermometer. While temperature is a perfectly well-defined physical quantity on the macro-scale, it is a meaningless concept on the level of a single atom or molecule. At the other extreme, we have one of the most complex (and still controversial) examples of emergence of human consciousness, which mysteriously emerges out of a caldron of interacting neurons. Consciousness cannot be found in any individual neuron, but is the collective property of the whole brain.

Example #1: Reynold's Boids

One of the most breathtakingly beautiful displays of nature is the synchronous fluid-like flocking of birds. It is also an excellent example of emergence in complex systems. Large or small, the magic of flocks is the very strong impression they convey of some intentional centralized control directing the overall traffic. Though ornithologists still do not have a complete explanation for this phenomenon, evidence strongly suggests that flocking is a decentralized activity, where each bird acts according to its local perceptions of what nearby birds are doing. Flocking is therefore a group behavior that emerges from collective action.

Craig Reynolds [264] programmed a set of artificial birds – which he called boids – so that each boid followed three simple local rules:
Land Warfare and Complexity, Part I: Mathematical Background and Technical Sourcebook

- **Rule 1**: maintain minimum distance from other objects (including other boids)
- **Rule 2**: match velocity of nearby boids
- **Rule 3**: move toward the perceived center of nearby boids

Each boid thus "sees" only what its neighbors are doing and acts accordingly. Reynolds found that the collective motion of all the boids was remarkably close to real flocking, despite the fact that there is nothing explicitly describing the flock as a whole. The boids initially move rapidly together to form a flock. The boids at the edges either slow down or speed up to maintain the flock's integrity. If the path bends or zigzags in any way, the boids all make whatever minute adjustments need to be made to maintain the group structure. If the path is strewn with obstacles, the boids flock around whatever is in their way naturally, sometimes temporarily splitting up to pass a an obstacle before reassembling beyond it. There is no central command that dictates this action.

The point of this example is not that the boids' behavior is a perfect replica of natural flocking – although it is a close enough match that Reynolds's model has attracted the attention of professional ornithologists – but that much of the boids' collective behavior is entirely unanticipated, and cannot be easily derived from the rules defining what each individual boid does.

**Example #2: Collective Decentralized Sorting**

Deneubourg, et. al. [213], have introduced a simple distributed sorting algorithm that is inspired by the self-organized way in which ant-colonies sort their brood.

Implemented by robot teams, their algorithm has the robots move about a fenced-in environment that is randomly littered with objects that can be scooped up. These robots (1) move randomly, (2) do not communicate with each other, (3) can perceive only those objects directly in front of them (but can distinguish between two or more types of objects with some degree of error), and (4) do not obey any centralized control. The probability that a robot picks up or puts down an object is a function of the number of the same objects that it has encountered in the past.

Coordinated by the positive feedback these simple rules induce between robots and their environment, the result, over time, is a seemingly intelligent, coordinated sorting activity. Clusters of randomly distributed objects spontaneously and quite naturally
emerge out a simple set of autonomous local actions having nothing at all to do with clustering per se; see figure 16.

The authors suggest that this system’s simplicity, flexibility, error tolerance and reliability compensate for their lower efficiency. One can argue, for example, that this collective sorting algorithm is much less efficient than a hierarchical one. The cost of having a hierarchy, though, is that the sorting would no longer be ant-like but would require a god-like oracle analyzing how many objects of what type are where, deciding how best to communicate strategy to the ants. Furthermore, the ants would require some sort of internal map, a rudimentary intelligence to deal with fluctuations and surprises in the environment (what if an object was not where the oracle said it would be?), and so on. In short, a hierarchy, while potentially more efficient, would of necessity have to be considerably more complex as well. The point Deneubourg, et. al. are making is that a much simpler collective decentralized system can lead to seemingly intelligent behavior while being more flexible, more tolerant of errors and more reliable that a hierarchical system.

Figure 16. Collective sorting by ant-like robots

Other examples of emergence include

- the characteristic spirals of the Belousov-Zhabotinski chemical reaction; see page 89

- the Navier-Stokes-like macroscopic behavior of a lattice gas that consists, on the micro-scale, of simple unit-bit billiards moving back and forth between discrete nodes along discrete links; see page 90

- globally ordered collective behavior in high-dimensional cellular automata systems that is locally featureless; see page 91

The macroscopic behavior in each of these examples is unexpected despite the fact that the details of the microscopic dynamics is well defined.
Edge-of-Chaos

One often hears the phrase *edge-of-chaos* in discussions of complex systems, as in "such and such a system appears poised at the edge-of-chaos." As this important concept is still a topic of some debate, we make a few comments regarding it.

Chris Langton [187] opens his "Life at the Edge of Chaos" paper at the *Artificial Life II* conference with the following intriguing question: "Under what conditions can we expect a dynamics of information to emerge spontaneously and come to dominate the behavior of a physical system?" While his question was, in that paper, motivated chiefly by an understanding that living organisms may be distinguished from inanimate matter by the fact that their behavior is clearly based on a complex dynamics of information, its roots extend considerably deeper.

Langton was able to provide a tentative answer to his question by examining the behavior of the entire rule space of elementary one-dimensional *cellular automata* rules (see discussion beginning on page 81) as parameterized by a single parameter $\lambda$. He found that as $\lambda$ is increased from its minimal to maximal values, a path is effectively traced in the rule space that progresses from fixed point behavior to simple periodicity to evolutions with longer and longer periods with increasing transients, passes through an intermediate transition region at a critical value $\lambda_c$, crosses over into a chaotic regime of steadily diminishing complexity until, eventually, the behavior is again completely predictable at the maximal value of $\lambda$ and complexity falls back to zero. Because the transition region represents the region of greatest *complexity* and lies between regions in which the behavior is either ordered or chaotic, Langton christened the transition region as the *edge-of-chaos*.

Langton's tentative answer to the question above is therefore: "We expect that information processing can emerge spontaneously and come to dominate the dynamics of a physical system in the vicinity of a critical phase transition." Langton speculates that the dynamics of phase transitions is fundamentally equivalent to the dynamics of information processing.

---

7 Elementary cellular automata are discrete dynamical systems. They consist of automata that live on sites of a one-dimensional lattice and that take on one of only two values – 0 or 1. Their dynamics is completely prescribed by a rule, $f$, that explicitly maps a state consisting of an automaton’s state and the states of the automaton’s left and right neighbors to either the value 0 or 1. Given a cellular automata rule $f$, Langton’s parameter $\lambda$ is defined to be the fraction of entries in the rule table for $f$ that get mapped to a non-zero value. For a more complete discussion of cellular automata, see page 81.
Figure 17. A Schematic illustration of the edge-of-chaos metaphor

Ordered Regime
perturbations die out

Complex Regime
poised to adapt and evolve

Chaotic Regime
effects of perturbations
propagate rapidly

Phase Transition

Strictly speaking, Langton's edge-of-chaos idea holds true only for the specific system in which it was discovered. Nonetheless, the idea has frequently been used as a general metaphor for the region in "complexity space" toward which complex adaptive systems appear to naturally evolve (see figure 17). Kauffman ([171], [172]), in particular, is a staunch advocate of the idea that systems poised at the edge-of-chaos are optimized, in some sense, to evolve, adapt and process information about their environment.

Effective computation, such as that required by life processes and the maintenance of evolvability and adaptability in complex systems, requires both the storage and transmission of information. If correlations between separated sites (or agents) of a system are too small – as they are in the ordered regime shown in figure 17 – the sites evolve essentially independently of one another and little or no transmission takes place. On the other hand, if the correlations are too strong – as they are the chaotic regime – distant sites may cooperate so strongly so as to effectively mimic each other's behavior, or worse yet, whatever ordered behavior is present may be overwhelmed by random noise; this, too, is not conducive to effective computation. It is only within the phase transition region – in the complex regime poised at the edge-of-chaos – that information can propagate freely over long distances without appreciable decay. However loosely defined, the behavior of a system in this region is best described as complex; i.e. it neither locks into an ordered pattern nor does it dissolve into an apparent randomness. Systems existing in this region are both stable enough to store information and dynamically amorphous enough to be able to successfully transmit it.

However intuitive the edge-of-chaos idea appears to be, one should be aware that it has received a fair amount of criticism in recent years. It is not clear, for example, how to even define complexity in more complicated systems like coevolutionary systems, much less imagine a phase transition between different complexity regimes. Even Langton's suggestion that effective computation within the limited domain of
cellular automata can take place only in the transition region has been challenged.  

**Complexity as a Measure?**

Related to the concept of the edge-of-chaos is the problem of determining what is meant by complexity, *as a measure*. That is to say, the problem of finding an objective measure by which an object X can be said to be more or less "complex" than object Y.

To set up the problem and in order to appreciate more fully the difficulty in quantifying complexity, consider figure 18. The figure shows three patterns: (1) an area of a regular two-dimensional Euclidean lattice, (2) a space-time view of the evolution of an elementary one-dimensional cellular automaton (see page 81 for discussion), and (3) a completely random collection of dots. These patterns illustrate the incongruity that exists between mathematically precise notions of entropy, or the amount of disorder in a system, and intuitive notions of complexity. Whereas pattern (2) is intuitively the most complex of the three patterns, it has neither the highest entropy (which belongs to pattern (3)) nor the lowest (which belongs to pattern (1)). Indeed, were we to plot our intuitive sense of complexity as a function of the amount of order or disorder in a system, it would probably look something like that shown in figure 19 (compare this figure to figure 17). The problem is to find an objective measure of the complexity of a system that matches this intuition.

Figure 18. Three patterns of varying "complexity"

We all have an intuitive feel for complexity. An oil painting by Picasso is obviously more "complex" than the random finger-paint doodles of a three-year-old. The works of Shakespeare are more "complex" than the rambling prose banged out on a typewriter by the proverbial band

---

of monkeys. Our intuition tells us that complexity is usually greatest in systems whose components are arranged in some intricate difficult-to-understand pattern or, in the case of a dynamical system, when the outcome of some process is difficult to predict from its initial state.

Figure 19. Complexity versus degree of order in a system

The problem is to articulate this intuition formally; to define a measure that not only captures our intuitive feel for what distinguishes the complex from the simple but also provides an objective basis for formulating conjectures and theories about complexity. While a universally accepted measure has yet to be defined (over 30 measures of complexity have been proposed in the research literature; see [86]), all such measures of complexity fall into two general classes:

- **Static Complexity**, which addresses the question of how an object or system is put together (i.e. only purely structural informational aspects of an object, or the patterns and/or strengths of interactions among its constituent parts), and is independent of the processes by which information is encoded and decoded.

- **Dynamic complexity**, which addresses the question of how much dynamical or computational effort is required to describe the information content of an object or state of a system.

Note that while a system's static complexity certainly influences its dynamical complexity, the two measures are clearly not equivalent. A system may be structurally rather simple (i.e. have a low static complexity), but have a complex dynamical behavior. (Think of the chaotic behavior of Feigenbaum's logistic equation, for example; see page 84).
What is the Difference Between *Chaos* and *Complexity*?

Very loosely speaking, it can be said that where *chaos* is the study of how simple systems can generate complicated behavior, *complexity* is the study of how complicated systems can generate simple behavior. Since both chaos and complex systems theory attempt to describe the behavior of dynamical systems, it should not be surprising to learn that both share many of the same tools, although, properly speaking, complex systems theory ought to be regarded as the superset of the two methodologies.

Table 5 lists behavioral characteristics of four basic kinds of dynamics: *ordered*, *random*, *chaotic* and *complex*.

**Table 5. A comparison among different types of dynamics**

<table>
<thead>
<tr>
<th>Examples</th>
<th>Ordered</th>
<th>Random</th>
<th>Chaotic</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictability</td>
<td>very high</td>
<td>none (statistical)</td>
<td>short times only</td>
<td>continually evolving</td>
</tr>
<tr>
<td>Effects of Small Perturbations</td>
<td>very small</td>
<td>none</td>
<td>exponential growth of initial errors</td>
<td>adaptation</td>
</tr>
<tr>
<td>Dimensionality (degrees of freedom)</td>
<td>finite</td>
<td>infinite</td>
<td>typically low</td>
<td>very high</td>
</tr>
<tr>
<td>Attractors</td>
<td>limit-points and limit-cycles</td>
<td>none</td>
<td>strange attractors</td>
<td>emergence vice attractors</td>
</tr>
<tr>
<td>Control</td>
<td>easy</td>
<td>hard</td>
<td>difficult (but ripe for exploitation)</td>
<td>self-adaptive</td>
</tr>
</tbody>
</table>

Cellular Automata

Cellular automata (CA) are a class of spatially and temporally discrete, deterministic mathematical systems characterized by local interaction and an inherently parallel form of evolution. First introduced by von Neumann in the early 1950s to act as simple models of biological self-reproduction, CA are prototypical models for complex systems and processes consisting of a large number of identical, simple, locally interacting components. The study of these systems has generated great interest over the years because of their ability to generate a rich spectrum of very complex patterns of behavior out of sets of relatively simple underlying rules. Moreover, they appear to capture many
essential features of complex self-organizing cooperative behavior observed in real systems.

Although much of the theoretical work with CA has been confined to mathematics and computer science, there have been numerous applications to physics, biology, chemistry, biochemistry, and geology, among other disciplines. Some specific examples of phenomena that have been modeled by CA include fluid and chemical turbulence, plant growth and the dendritic growth of crystals, ecological theory, DNA evolution, the propagation of infectious diseases, urban social dynamics, forest fires, and patterns of electrical activity in neural networks. CA have also been used as discrete versions of partial differential equations in one or more spatial variables. They have most recently been used to simulate some aspects of military combat [323].

The best sources of information on CA are conference proceedings and collections of papers, such as the one's edited by Bocca (29), Gutowitz (120), Preston (254) and Wolfram (321)-[322]. An excellent review of how CA can be used to model physical systems is given by Toffoli and Margolus [304].

While there is an enormous variety of particular CA models – each carefully tailored to fit the requirements of a specific system – most CA models usually possess these five generic characteristics:

- **discrete lattice of cells**: the system substrate consists of a one-, two- or three-dimensional lattice of cells
- **homogeneity**: all cells are equivalent
- **discrete states**: each cell takes on one of a finite number of possible discrete states
- **local interactions**: each cell interacts only with cells that are in its local neighborhood (figure 23 shows some common neighborhoods in two dimensions)
- **discrete dynamics**: at each discrete unit time, each cell updates its current state according to a transition rule taking into account the states of cells in its neighborhood

**Example #1: One-dimensional CA**

For a one-dimensional CA, the value of the ith cell at time t – denoted by $c_i(t)$ – evolves in time according to a "rule" $F$ that is a function of $c_i(t)$ and other cells that are within a range $r$ (on the left and right) of $c_i(t)$:
\[ c_i(t) = F[c_{i-r}(t-1), c_{i-r+1}(t-1), \ldots, c_{i+r-1}(t-1), c_{i+r}(t-1)]. \]

Since each cell takes on one of \( k \) possible values – that is, \( c_i(t) \in \{0,1,2,\ldots,k\} \) – the rule \( F \) is completely defined by specifying the value assigned to each of the \( k^{2r+1} \) possible \((2r+1)\)-tuple configurations for a given range-\( r \) neighborhood:

<table>
<thead>
<tr>
<th>( c_{i-r}(t-1) )</th>
<th>( \ldots )</th>
<th>( c_i(t-1) )</th>
<th>( \ldots )</th>
<th>( c_{i+r}(t-1) )</th>
<th>( c_i(t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( F(0,0,\ldots,0) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>( F(0,0,\ldots,1) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>( k )</td>
<td>( k )</td>
<td>( k )</td>
<td>( F(k,k,\ldots,k) )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since \( F \) itself assigns any of \( k \) values to each of the \( k^{2r+1} \) possible \((2r+1)\)-tuples, the total number of possible rules is an exponentially increasing function of both \( k \) and \( r \). For the simplest case of nearest neighbors (range \( r=1 \)) and \( k=2 \) (\( c_i = 0 \) or \( 1 \), for example, there are \( 2^6=256 \) possible rules. Increasing the number of values each cell can take on to \( k=3 \) (but keeping the radius at \( r=1 \)) increases the rule-space size to \( 3^3 \approx 7 \times 10^{12} \).

Figure 20 shows the time evolution of a nearest-neighbor (radius \( r=1 \)) rule where \( c \) is equal to either 0 or 1. The row of eight boxes at the top of the figure shows the explicit rule-set, where – for visual clarity – a box has been arbitrarily colored "black" if the value \( c=1 \) and "white" if \( c=0 \). For each combination of three adjacent cells in generation 0, the rule \( F \) assigns a particular value to the next-generation center cell of the triplet. Beginning from an initial state (at time=0) consisting of the value zero everywhere except the center site, that is assigned the value 1, \( F \) is applied synchronously at each successive time step to each cell of the lattice. Each generation is represented by a row of cells and time is oriented downwards. The first image shows a blowup of the first five generations of the evolution. The second shows 300 generations. The figure illustrates the fact that simple rules can generate considerable complexity.

The space-time pattern generated from a single nonzero cell by this particular rule has a number of interesting properties. For example, it consists of a curious mixture of ordered behavior along the left-hand-side and what appears to be disordered behavior along the right-hand-side, separated by a corrugated boundary moving towards the left at a "speed" of about 1/4 cells per "clock" tick. In fact, it can be shown that, despite starting from an obviously non-random initial state and evolving according to a fixed deterministic rule, the temporal sequence of vertical values is completely random. Systems having the ability to deterministically generate randomness from non-random input are called autoplectic systems.
Figure 20. Example of a one-dimensional CA

As another example, consider the rule shown at the top of figure 21. Its space-time evolution, starting from a random initial state, is shown at the bottom of the figure. Note that this space-time pattern can be described on two different levels: either on the cell-level, by explicitly reading off the values of the individual cells, or on a higher-level by describing it as a sea of particle-like structures superimposed on a periodic background. In fact, following a small initial transient period, temporal sections of this space-time pattern are always of the form "...BBBPBBB...BB...BBBPBBB...BBBPBBB...BB...BBBPBB...BBB...BBBPBBB...BBB...BBBPBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBBPBBB...BBB...BBP", where "B" is a state of the periodic background consisting of repetitions of the sequence "1001101111000" (with spatial period 14 and temporal period 7), and the P's represent "particles." The particle pattern $P = "11111000"$, for example, repeats every four steps while being displaced two cells to the
left; the particle \( P = \text{"11101011000"} \) repeats every ten steps while being displaced two cells to the right.

Figure 21. Evolution of a one-dimensional CA starting from a random initial state

\[
\begin{array}{c|c}
\text{■} & \text{□} \\
\hline
\text{□} & \text{□} \\
\text{□} & \text{□} \\
\text{□} & \text{□} \\
\text{□} & \text{□} \\
\text{□} & \text{□} \\
\text{□} & \text{□} \\
\text{□} & \text{□} \\
\text{□} & \text{□} \\
\end{array}
\]

Rule

Although the underlying dynamics describing this system is very simple, and entirely deterministic, there is an enormous variety, and complexity, of emergent particle-particle interactions. Such simple systems are powerful reminders that complex higher-level dynamics need not have a complex underlying origin. Indeed, suppose that we had been shown such a space-time pattern but were told \textit{nothing whatsoever about its origin}. How would we make sense of its dynamics? Perhaps the only reasonable course of action would be to follow the lead of any good experimental particle-physicist and begin cataloging the various possible particle states and interactions: \textit{there are} \( N \) \textit{particles of size} \( s \) \textit{moving to the left with speed} \( v \), \textit{when a particle} \( p \) \textit{of type} \( P \) \textit{collides with} \( q \) \textit{of type} \( Q \), \textit{the result is the set of particles} \( \{p, \ldots, q\} \); and so on. It would take a tremendous leap of intuition to fathom the utter simplicity of the real dynamics.

In general, the behavior of CA is strongly reminiscent of the kinds of behavior observed in continuum dynamical systems, with simple rules yielding steady-state behaviors consisting of fixed points or limit cycles, and complex rules giving rise to behaviors that are analogous to deterministic chaos. In fact, there is extensive empirical evidence suggesting that patterns generated by all (one-dimensional) CA
evolving from disordered initial states fall into one of only four basic behavioral classes:

- **Class 1:** evolution leads to a homogenous state, in which all cells eventually attain the same value

- **Class 2:** evolution leads to either simple stable states or periodic and separated structures

- **Class 3:** evolution leads to chaotic nonperiodic patterns

- **Class 4:** evolution leads to complex, localized propagating structures

All CA within a given class yield qualitatively similar behavior. While the behaviors of rules belonging to the first three rule classes bear a strong resemblance to those observed in continuous systems – the homogenous states of class 1 rules, for example, are analogous to fixed-point attracting states in continuous systems, the asymptotically periodic states of class 2 rules are analogous to continuous limit cycles and the chaotic states of class 3 rules are analogous to strange attractors – the more complicated localized structures emerging from class 4 rules do not appear to have any obvious continuous analogues (although such structures are well characterized as being soliton-like in their appearance).

Figure 22 shows a few examples of the kinds of space-time patterns generated by binary (k=2) nearest-neighbor (r=1) in one dimension and starting from random initial states.

Figure 23 shows examples of some commonly used neighborhood structures in two dimensions. These include (1) the von Neumann neighborhood, which consists of the four cells that are horizontally and vertically adjacent to the center cell, (2) the Moore neighborhood, that consists of all eight nearest-neighbor cells on a two-dimensional Euclidean lattice, and (3) the Hexagonal neighborhood, that consists of all nearest-neighbor cells on a hexagonal lattice.
Figure 22. Space-time evolution of nine different nearest neighbor one-dimensional CA starting from random initial states

Figure 23. Examples of CA neighborhoods in two dimensions

- **von Neumann**
- **Moore**
- **Hexagonal**

**Example #2: Conway's Life**

"It's probable, given a large enough Life space, initially in a random state, that after a long time, intelligent self-reproducing animals will emerge and populate some parts of the space." – John H. Conway

Perhaps the most widely known CA is the game of Life, invented by John H. Conway, and popularized extensively by Martin Gardner in his "Mathematical Games" department in Scientific American in the early 1970s.

Life is "played" using the 9-neighbor Moore neighborhood (see figure 23), and consists of (1) seeding a lattice with some pattern of "live" and "dead" cells, and (2) simultaneously (and repeatedly) applying the following three rules to each cell of the lattice at discrete time steps:
\textbullet \textit{Birth:} replace a previously dead cell with a live one if exactly 3 of its neighbors are alive

\textbullet \textit{Death:} replace a previously live cell with a dead one if either (1) the living cell has no more than one live neighbor (i.e. it dies of isolation), or (2) the living cell has more than three neighbors (i.e. it dies of overcrowding)

\textbullet \textit{Survival:} retain living cells if they have either 2 or 3 neighbors

One of the most intriguing patterns in Life is an oscillatory propagating pattern known as the "glider." Shown on the left-hand-side of figure 24, it consists of 5 "live" cells and reproduces itself in a diagonally displaced position once every four iterations. When the states of Life are projected onto a screen in quick succession by a fast computer, the glider gives the appearance of "walking" across the screen. The propagation of this pseudo-stable structure can also be seen as a self-organized emergent property of the system. The right-hand-side of figure 24 shows a still-frame in the evolution of a pattern known as a "glider-gun," which shoots-out a glider once every 30 iteration steps.

What is remarkable about this very simple appearing rule is that one can show that it is capable of universal computation. This means that with a proper selection of initial conditions (i.e. the initial distribution of "live" and "dead" cells), Life can be turned into a general purpose computer. This fact fundamentally limits the overall predictability of Life's behavior.

\textbf{Figure 24. Glider patterns in Conway's Life}

The well known Halting Theorem, for example, asserts that there cannot exist a general algorithm for predicting when a computer will halt its execution of a given program [107]. Given that Life is a universal computer – so that the Halting Theorem applies – this means that one cannot, in general, predict whether a particular starting configuration of live and dead cells will eventually die out. No shortcut
is possible, even in principle. The best one can do is to sit back and patiently await Life's own final outcome.

Put another way, this means that if you want to predict Life's long-term behavior with another "model" or by using, say, a partial differential equation, you are doomed to fail from the outset because its long-term behavior is effectively unpredictable. Life - like all computationally universal systems - defines the most efficient simulation of its own behavior.

**Example #3: Belousov-Zhabotinski Reaction**

The Belousov-Zhabotinski (BZ) reaction is a chemical reaction consisting of simple organic molecules that is characterized by spectacular oscillating temporal and spatial patterns. One variant of the BZ reaction involves the reaction of bromate ions with an organic substrate (typically malonic acid) in a sulfuric acid solution with cerium (or some other metal-ion catalyst). When this mixture is allowed to react exothermally at room temperature, interesting temporal and spatial oscillations (i.e. chemical waves) result. The system oscillates, changing from yellow to colorless and back to yellow about twice a minute, with the oscillations typically lasting for over an hour (until the organic substrate is exhausted).

These patterns are an example of what are sometimes called dissipative structures, which arise in many complex systems. Dissipative structures are dynamical patterns that retain their organized state by persistently dissipating matter and energy into an otherwise thermodynamically open environment.

![Figure 25. Example of self-organization in a two-dimensional CA](image)

Figure 25 shows a sample evolution of a CA model of this reaction, in which cells are identified with the reacting molecules, and are colored "black" if they are "active" and "white" if they are "inactive," according to the reaction rules. The spatial and temporal patterns that emerge...
from the initially random mixture of states are also a good general example of how CA can be used to model self-organization.

**Example #4: Lattice Gases**

Lattice gases are micro-level rule-based simulations of macro-level fluid behavior. Lattice-gas models provide a powerful new tool in modeling real fluid behavior. The idea is to reproduce the desired macroscopic behavior of a fluid by modeling the underlying microscopic dynamics.

It can be shown that three basic ingredients are required to achieve an emergence of a suitable macrodynamics out of a discrete microscopic substrate: (1) local thermodynamic equilibrium, (2) conservation laws, and (3) a "scale separation" between the levels at which the microscopic dynamics takes place (among kinetic variables living on a micro-lattice) and the collective motion itself appears (defined by hydrodynamical variable on a macro-lattice). Another critical feature is the symmetry of the underlying lattice.

While there are many variants of the basic model, one can show that there is a well-defined minimal set of rules that define a lattice-gas system whose macroscopic behavior reproduces that predicted by the Navier-Stokes equations\(^9\) exactly. In other words, there is critical "threshold" of rule size and type that must be met before the continuum fluid behavior is matched, and once that threshold is reached the efficacy of the rule-set is no longer appreciably altered by additional rules respecting the required conservation laws and symmetries.

**Figure 26. Two-dimensional lattice-gas simulation of a fluid**

![Time snapshots](image)

\(\text{time} = 0\) \hspace{1cm} \(\text{time} = 200\) \hspace{1cm} \(\text{time} = 500\)

Figure 26 shows a few snapshots of the evolution of a two-dimensional lattice gas starting from an initial condition in which there is a tightly packed region of particles at the center of the lattice. Notice how this central region expands rapidly outward, and is very reminiscent of the

---

\(^9\) The Navier-Stokes equations are a set of analytically intractable coupled nonlinear partial differential equations describing fluid flow.
effect a dropped stone has on an initially stagnant pool of water. The most striking feature is the circular sound wave, which is circular despite the fact that the microscopic dynamics takes place on a square lattice. The lattice gas "rules" thus force a symmetry that is not present in the microscopic dynamics to emerge on the macro-scale.

**Example #5: Collective Behavior in Higher Dimensions**

Chate and Manneville\(^ {10} \) have examined a wide variety of cellular automata that live in dimensions four, five and higher. They found many interesting rules that while being essentially featureless locally, nonetheless show a remarkably ordered global behavior.

Figure 27, for example, plots the probability that a cell has value 1 at time \( t+1 \) – labeled \( P_{t+1} \) – versus the probability that a cell had value 1 at time \( t \) – labeled \( P_t \) – for a particular four dimensional cellular automaton rule. The rule itself is unimportant, as there are many rules that display essentially the same kind of behavior. The point is that while the behavior of this rule is locally featureless – its space-time diagram would look like static on a television screen – the global density of cells with value 1 jumps around in quasi-periodic fashion. We emphasize that this quasi-periodicity is a global property of the system, and that no evidence for this kind of behavior is apparent in the local dynamics.

Here again we see a pristine example of the three basic elements of emergence: (1) the global phenomenon (in this case the cell-value density) emerges out of an interaction of a large number of simple components (lattice cells of a cellular automaton), (2) there is no evidence of the global phenomenon on the local level, and (3) the global phenomenon obeys a separate dynamics (in this case, quasi-periodicity).

Other Variants

There are as many different variants of the basic CA algorithm as there are ways of generalizing the five fundamental characteristics of what makes up a CA system. Here are a few:

- **Probabilistic CA** (PCA). Probabilistic CA are cellular automata in which the deterministic state-transitions are replaced with specifications of the probabilities of the cell-value assignments. Since such systems have much in common with certain statistical mechanical models, analysis tools from physics are often borrowed for their study.

- **Non-homogeneous CA**. These are CA in which the state-transition rules are allowed to vary from cell to cell. The simplest such example is one where there are only two different rules randomly distributed throughout the lattice. Kauffman [171] has studied the other extreme in which the lattice is randomly populated with all $2^k$ possible Boolean functions of $k$ inputs.

- **Coupled-map Lattices**. These are models in which continuity is restored to the state space. That is to say, the cell values are no longer constrained to take on only the values 0 and 1 as in the examples discussed above, but can now take on arbitrary real values. First introduced by Kaneko [167], such systems are simpler than partial differential equations but more complex than generic CA.

Genetic Algorithms

Genetic algorithms (GAs) are a class of heuristic search methods and computational models of adaptation and evolution based on natural selection.

In nature, the search for beneficial adaptations to a continually changing environment (i.e. evolution) is fostered by the cumulative evolutionary knowledge that each species possesses of its forebears. This knowledge, which is encoded in the chromosomes of each
member of a species, is passed from one generation to the next by a mating process in which the chromosomes of "parents" produce "offspring" chromosomes.

GAs mimic and exploit the genetic dynamics underlying natural evolution to search for optimal solutions of general combinatorial optimization problems. They have been applied to the Traveling Salesman Problem, VLSI circuit layout, gas pipeline control, the parametric design of aircraft, neural net architecture, models of international security, and strategy formulation.

While their modern form is derived mainly from John Holland's work in the 1960s [142], many key ideas such as using "selection of the fittest" like population-based selection schemes and using binary strings as computational analogs of biological chromosomes, actually date back to the late 1950s. More recent work is discussed by Goldberg [113], Davis [65] and Michalewicz [215] and in conference proceedings edited by Forrest [99]. A comprehensive review of the current state-of-the-art in genetic algorithms is given by Mitchell [220].

The basic idea behind GAs is very simple. Given a "problem" – which can be as well-defined as maximizing a function over some specified interval or as seemingly ill-defined and open-ended as evolution itself, where there is no a-priori discernible or fixed function to either maximize or minimize – GAs provide a mechanism by which the solution space to that problem is searched for "good solutions." Possible solutions are encoded as chromosomes (or, sometimes, as sets of chromosomes), and the GA evolves one population of chromosomes into another according to their fitness by using some combination (and/ or variation) of the genetic operators of crossover and mutation. A solution search space together with a fitness function is called a fitness landscape. Eventually, after many generations, the population will, in theory, be composed only of those chromosomes whose fitness values are clustered around the global maximum of the fitness landscape.

**Genetic Operators**

Each chromosome is usually defined to be a bit-string, where each bit position (or "locus") takes on one of two possible values (or "alleles"), and can be imagined as representing a single point in the "solution space." The fitness of a chromosome effectively measures how "good" a solution that chromosome represents to the given problem. Aside from its intentional biological roots and flavoring, GAs can be thought of as parallel equivalents of more conventional serial optimization techniques: rather than testing one possible solution after another, or moving from point to point in the solution phase-space, GAs move from entire populations of points to new populations.
Figure 28 shows examples of the three basic genetic operations of reproduction, crossover and mutation, as applied to a population of 8-bit chromosomes. Reproduction makes a set of identical copies of a given chromosome, where the number of copies depends on the chromosome's fitness. The crossover operator exchanges subparts of two chromosomes, where the position of the crossover is randomly selected, and is thus a crude facsimile of biological sexual recombination between two single-chromosome organisms. The mutation operator randomly flips one or more bits in the chromosome, where the bit positions are randomly chosen. The mutation rate is usually chosen to be small.

While reproduction generally rewards high fitness, and crossover generates new chromosomes whose parts, at least, come from chromosomes with relatively high fitness (this does not guarantee, of course, that the crossover-formed chromosomes will also have high fitness; see below), mutation seems necessary to prevent the loss of diversity at a given bit-position. For example, were it not for mutation, a population might evolve to a state where the first bit-position of each chromosome contains the value 1, with there being no chance of reproduction or crossover ever replacing it with a 0.

Figure 28. Schematic representation of the basic GA operators

| Reproduction: | \[00101110\] \[00101110\] \[00101110\] |
| Crossover:    | \[01100101\] \[10100111\] |
|               | \[01100101\] \[10100001\] |
| Mutation:     | \[00101110\] \[00101010\] |

**The Basic GA Recipe**

Although GAs, like CA, come in many different flavors, and are usually fine-tuned in some way to reflect the nuances of a particular problem, they are all more or less variations of the following basic steps (see figure 29):
**Step 1:** begin with a randomly generated population of chromosome-encoded "solutions" to a given problem

**Step 2:** calculate the fitness of each chromosome, where fitness is a measure of how well a member of the population performs at solving the problem

**Step 3:** retain only the fittest members and discard the least fit members

**Step 4:** generate a new population of chromosomes from the remaining members of the old population by applying the operations reproduction, crossover, and mutation (see figure 28)

**Step 5:** calculate the fitness of these new members of the population, retain the fittest, discard the least fit, and re-iterate the process

---

**Example #1: Function Maximization**

As a concrete example, suppose our problem is to maximize the fitness function \( f(x) = x^2 \), using six 6-bit chromosomes of the form \( C = (c_1, c_2, ..., c_6) \), where each \( c_i \) is equal to either 0 or 1. C's fitness, \( f(C) \), is determined by first converting its binary representation into a base-10 equivalent value and squaring: 
\[
f(C) = (c_1 + 2c_2 + 2^2c_3 + 2^3c_4 + 2^4c_5 + 2^5c_6)^2.
\]

The first step is to construct six random bit-strings representing the initial population:
Land Warfare and Complexity, Part I: Mathematical Background and Technical Sourcebook

<table>
<thead>
<tr>
<th>C_1 = (101101)</th>
<th>C_2 = (010110)</th>
<th>C_3 = (111001)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_4 = (101011)</td>
<td>C_5 = (010001)</td>
<td>C_6 = (011101)</td>
</tr>
</tbody>
</table>

These chromosomes have fitness values of 2025, 484, 3249, 1849, 289 and 841, respectively. The average fitness is 1456. By luck of the fitness-scaled draw, where the number of copies of a given chromosome is determined according to its fitness, scaled by the average fitness of the entire population, three copies of C_5 are made for the next population (owing to its relatively high fitness), one copy each for chromosomes C_1, C_4 and C_6 and none for the remaining chromosomes. These copies form the mating population.

Next, we randomly pair up the new chromosomes, and perform the genetic crossover operation at randomly selected bit-positions - chromosomes C_1 and C_4 exchange their last three bits, C_2 and C_6 exchange their last four bits, and C_3 and C_5 exchange their last bit:

| C_1 exchange with C_4 at bit 3:              | (101.101) x (111.001) | → | (101001) |
| C_2 exchange with C_5 at bit 2:              | (111.001) x (01.1101) | → | (111011) |
| C_3 exchange with C_6 at bit 5:              | (111001.1) x (101011.1) | → | (111001) |
| C_4 exchange with C_1 at bit 3:              | (111001) x (101.101)  | → | (111001) |
| C_5 exchange with C_2 at 5:                  | (111001.1) x (111001.1) | → | (111001) |
| C_6 exchange with C_3 at bit 2:              | (01.1101) x (11.1001) | → | (011001) |

Finally, we mutate each bit of the resulting chromosomes with some small probability - say \( p_{\text{mutation}} = 0.05 \). In our example we find that values of the 5th bit in C_4 and 6th bit in C_6 are flipped. The resulting strings make up our 2nd generation chromosome population. By chance, the first loop through the algorithm has successfully turned up the most-fit chromosome - C_5 = (111111) \( \rightarrow f(C_5) = 63^5 = 3969 \) - but in general the entire procedure would have to be repeated many times to approach the "desired" solution.

The table below summarizes the above steps:

<table>
<thead>
<tr>
<th>Initial Population</th>
<th>Initial Fitness</th>
<th>Expected Copies</th>
<th>Actual Copies</th>
<th>Mating Population</th>
<th>Crossover Operation(^1)</th>
<th>Mutation Operation</th>
<th>New Fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>(101101)</td>
<td>2025</td>
<td>1.4</td>
<td>1</td>
<td>(101101)</td>
<td>(134) ( \rightarrow ) (101001)</td>
<td>(101001)</td>
<td>1681</td>
</tr>
<tr>
<td>(010110)</td>
<td>484</td>
<td>0.3</td>
<td>0</td>
<td>(111001)</td>
<td>(226) ( \rightarrow ) (111101)</td>
<td>(111001)</td>
<td>3481</td>
</tr>
<tr>
<td>(111001)</td>
<td>3249</td>
<td>2.2</td>
<td>3</td>
<td>(111001)</td>
<td>(355) ( \rightarrow ) (111001)</td>
<td>(111001)</td>
<td>3249</td>
</tr>
<tr>
<td>(101011)</td>
<td>1849</td>
<td>1.3</td>
<td>1</td>
<td>(111001)</td>
<td>(431) ( \rightarrow ) (111101)</td>
<td>(111011)</td>
<td>3969</td>
</tr>
<tr>
<td>(010001)</td>
<td>289</td>
<td>0.2</td>
<td>0</td>
<td>(101011)</td>
<td>(558) ( \rightarrow ) (101011)</td>
<td>(101010)</td>
<td>1764</td>
</tr>
<tr>
<td>(011101)</td>
<td>841</td>
<td>0.6</td>
<td>1</td>
<td>(011101)</td>
<td>(622) ( \rightarrow ) (010101)</td>
<td>(011001)</td>
<td>625</td>
</tr>
</tbody>
</table>

\(^1\) The crossover operator \((xyz)\) means that chromosomes C_x and C_y exchange bits at the z\(^{th}\) bit.
Example #2: Local Forecasting of High-Dimensional Chaotic Dynamics

The first example of using a genetic algorithm to maximize the value of the function \( f(x) = x^3 \), discussed above, is deliberately simple and was chosen mainly for its pedagogical value. The second example, discussed below, shows more of the real power of genetic algorithms.

Meyer and Packard\(^ {11} \) use a genetic algorithm to learn patterns in data produced by a high-dimensional chaotic attractor. The patterns are relationships between a region of the attractor and the future behavior of chaotic orbits that pass through this region. They find that a genetic algorithm gives accurate local profiles of the attractor that provide forecasts of behavior.

Meyer and Packard first form a finite set of \( N \) data points \( \{\xi_1,\xi_2,...,\xi_N\} \) from a continuous variable representing the evolution of a high-dimensional chaotic system. From this set of points, they construct a sequence of points in a corresponding \( d \)-dimensional space \( X = \{\tilde{x}_d,\tilde{x}_{d+1},...,\tilde{x}_N\} \) where each \( \tilde{x}_t = (\xi_{t-d+1},...,\xi_t) \) is a set of past values in the time series. The problem is to determine a map from the past values \( (\xi_{t-d+1},...,\xi_t) \) to the future value at time \( t', \xi_{t'}, \) or, equivalently, to find a map from \( \tilde{x}_t \) to \( y_t = \xi_{t'} \).

Meyer and Packard search for specific patterns of the form

\[ (\alpha_i < \xi_i < \beta_i) \land (\alpha_j < \xi_j < \beta_j) \land \cdots \land (\alpha_k < \xi_k < \beta_k), \]

where "\( \land \)" means AND. In other words, they seek intervals that each of the past values of the set must be in to predict the future value of the trajectory at a given later point to some arbitrarily selected threshold error. The genetic algorithm is used to search for the optimal patterns of this form, mutations adding or deleting conditions to an existing "candidate" conjunction, crossover taking two existing candidates and exchanging roughly half of the constraints of one with the other.

Figure 30 shows the four best predictive patterns for the chaotic system chosen for this analysis. In the figure, the set \( X \), appearing on the left hand side, represents a set of 30 points from which the value of \( y^* \), shown at the right, is to be predicted. The intervening 150 points, shown in grey, represent intermediate times during which the chaotic system is allowed to evolve; no information is extracted from the data during this interval. The value of \( y^* \) is to be predicted solely from the 30 values contained in the set \( X \).

\(^{11} \) Thomas P. Meyer and N. H. Packard, "Local forecasting of high-dimensional chaotic dynamics," pages 249-263 in reference [44].
Notice the intermediate divergence of each of the trajectories and how they all subsequently collapse to a narrow range of values around the desired y'. In order to fully appreciate how impressive a find the predictive patterns shown in figure 30 really are, keep in mind the extraordinarily vast space of possible patterns that the genetic algorithm is asked to search through. Even if the each of 30 x values are allowed to take on one of only two values – remember that, in fact, they can take on a continuum of values – there are still $2^{30} \approx 10^9$ possible conjunctive patterns to search through!

Figure 30. The four best patterns in X yielding the value y*, as found by a genetic algorithm

\[ \text{a) } \]

\[ \text{b) } \]

\[ \text{c) } \]

\[ \text{d) } \]

**Other Variants**

There are several different variants of the basic genetic algorithm as outlined above:

- **Classifier Systems.** Classifier systems were introduced by John Holland as an attempt to apply genetic algorithms to cognitive tasks. They are similar to production systems of the "if...then" variety in artificial intelligence. A classifier system typically consists of (1) a set of detectors (or input devices) that provide information to the system about the state of the external environment, (2) a set of effectors (or output devices) that transmit the classifier's conclusions to the external environment,
(3) a set of rules (or classifiers), consisting of a condition and action, and (4) a list of messages. Learning is supervised as in multilayered neural networks (see page 116).

- **Evolutionary Programming.** Evolutionary programming is an early variant of genetic algorithms and is mainly distinguished from the conventional genetic algorithm by not incorporating crossover as an operator.

- **Genetic Programming.** Genetic programming is essentially an application of genetic algorithms to computer programs. Typically the genome is represented by a LISP expression, so that what evolves is a population of programs, rather than bit-strings as in the case of a usual genetic algorithm. For references see Koza [179] and the WWW sources listed in the appendix.

### Self-Organized Criticality

Self-organized criticality (SOC) describes a large body of both phenomenological and theoretical work having to do with a particular class of time-scale invariant and spatial-scale invariant phenomena. As with many of the terms and concepts associated with nonlinear dynamics and complex systems, its meaning has been somewhat diluted and made imprecise since its introduction a few years ago, in large part due to the veritable explosion of articles on complex systems appearing in the popular literature. Fundamentally, SOC embodies the idea that dynamical systems with many degrees of freedom naturally self-organize into a critical state in which the same events that brought that critical state into being can occur in all sizes, with the sizes being distributed according to a power-law.

"Criticality" here refers to a concept borrowed from thermodynamics. Thermodynamic systems generally get more ordered as the temperature is lowered, with more and more structure emerging as cohesion wins over thermal motion. Thermodynamic systems can exist in a variety of phases – gas, liquid, solid, crystal, plasma, etc. – and are said to be critical if poised at a phase transition. Many phase transitions have a critical point associated with them, that separates one or more phases. As a thermodynamic system approaches a critical point, large structural fluctuations appear despite the fact the system is driven only by local interactions. The disappearance of a characteristic length scale in a system at its critical point, induced by these structural fluctuations, is a characteristic feature of thermodynamic critical phenomena and is universal in the sense that it is independent of the details of the system's dynamics.
The kinds of structures SOC seeks to describe the underlying mechanisms for look like equilibrium systems near critical points but are not near equilibrium; instead, they continue interacting with their environment, "tuning themselves" to a point at which critical-like behavior appears. In contrast, thermodynamic phase transitions usually take place under conditions of thermal equilibrium, where an external control parameter such as temperature is used to tune the system into a critical state.

Introduced in 1988 by Bak, Chen and Wiesendfeld [15], SOC is arguably the only existing holistic mathematical theory of self-organization in complex systems, describing the behavior of many real systems in physics, biology and economics. It is also a universal theory in that it predicts that the global properties of complex systems are independent of the microscopic details of their structure, and is therefore consistent with the "the whole is greater than the sum of its parts" approach to complex systems. Put in the simplest possible terms, SOC asserts that complexity is criticality. That is to say, that SOC is nature's way of driving everything towards a state of maximum complexity.

In general, SOC appears to be prevalent in systems that have the following properties:

- they have many degrees of freedom
- their parts undergo strong local interactions
- the number of parts is usually conserved
- they are driven by being slowly supplied with "energy" from an exogenous source
- energy is rapidly dissipated within the system

In systems that have these properties, SOC itself is characterized by

- a self-organized drive towards the critical state
- intermittently triggered ("avalanche"-style) release of energy in the critical state
- sensitivity to initial conditions (i.e. the trigger can be very small)\textsuperscript{12}

\textsuperscript{12} Sensitivity to initial conditions is usually a trademark of chaos in dynamical systems. Unlike fully chaotic systems, however, in which nearby trajectories diverge exponentially, the distance between two trajectories in
the critical state is maintained without any external "tuning"

These ideas will be explained more fully in the example that follows.

**Example: Sandpiles**

To better illustrate the concept of SOC, consider a "toy model" of avalanches: A mechanical arm holds a large quantity of sand and sits securely in place some distance above a flat circular table. Slowly – individual grain by individual grain – the arm releases its store of sand. The sand thus begins forming a pile beneath the arm.

At first, the grains all stay relatively close together near where they all fall onto the pile. Then they begin piling up on top of one another, creating a pile with a small slope. Every once and a while the slope becomes too steep somewhere, and a few grains slide down in a small avalanche. As the mechanical arm continues dispensing grains of sand, the average slope of the pile of sand beneath it steepens, and the average size of the resulting avalanches increases. The size of the pile stops growing when the amount of sand added to the pile is balanced by the amount of sand that falls off the circular table. This state is the critical state.

What is special about the critical state is that when a grain of sand is added to a sandpile in this state, it can spawn an avalanche of any size, from the smallest avalanche consisting of only a few grains to a major "catastrophe" involving very many grains to no avalanche at all. Moreover, the size of an avalanche does not depend on the grain of sand that triggers it. However, the frequency f of avalanches of a size greater than or equal to a given size s is related to s by a power-law: \( f \propto 1/s^{-\beta} \), for some \( \beta > 0 \); a relationship that, according to Bak, et. al., is the signature characteristic of SOC (see figure 31). There is thus no such thing as an avalanche of average size. An estimate only gets larger as more and more avalanches are averaged together. The critical state is also stable: because even the largest avalanches involve only a small fraction of the total number of grains in the sandpile, once a pile has evolved to its critical state, it stays poised close to that state forever.

There is strong evidence to suggest that just as sandpiles self-organize into a critical state, so do many real complex systems naturally evolve, or "tune themselves," to a critical state, in which a minor event can, via systems undergoing SOC grows at a much slower (power-law) rate. Systems undergoing SOC are therefore only "weakly chaotic." There is an important difference between fully developed chaos and weak chaos: fully developed chaotic systems have a characteristic time scale beyond which it is impossible to make predictions about their behavior; no such time scale exists for weakly chaotic systems, so that long-time predictions may be possible.
a cascading series of chain-reactions, involve any number of elements of the system.

Figure 31. Power-law distribution from a computer simulation of a two-dimensional sandpile cellular automaton

The critical state is an attractor for the dynamics: systems are inexorably driven toward it for a wide variety of initial conditions. Frequently cited examples of SOC include the distribution of earthquake sizes, the magnitude of river flooding, and the distribution of solar flare x-ray bursts, among others. Conway’s Life-game CA-rule (see page 87), which is a crude model of social interaction, appears to self-organize to a critical state when driven by random mutations. Another vivid example of SOC is the extinction of species in natural ecologies. In the critical state, individual species interact to form a coherent whole, poised in a state far out of equilibrium. Even the smallest disturbances in the ecology can thus cause species to become extinct. Real data show that there are typically many small extinction events and few large ones, though the relationship does not quite follow the same linear power-law as it does for avalanches. Bak and Chen [15] have also speculated that “throughout history, wars and peaceful interactions might have left the world in a critical state in which conflicts and social unrest spread like avalanches.”

---

Fractals

Another characteristic feature of many complex systems is some form of a fractal structure. Just as structural fluctuations near phase transitions have no characteristic scales, self-similar fractal structures appear on all size scales and thus possess no characteristic length scales. Familiar examples of fractals include fractal coastlines, mountain landscapes and cloud formations. While fractals may be ubiquitous in nature, however, the underlying dynamical mechanisms are far from clear. It is reasonable to speculate that a common mechanism may be found using notions of thermodynamic criticality. Bak, et. al., suggest that fractal structures are the "spatial fingerprints" of SOC.

Consider a fractal time series. One of its key features is that it cannot be reduced to a series of periodic signals plus a noise term of the form \( x(t) = x_0 \sin(ft) + \text{Noise}(t) \). If a fractal time series could be expressed in this form, the contribution due to the noise term would average out as \( t \to \infty \) and the signal would have a well-defined average value for its frequency. Instead, a fractal time series is characterized by a distribution of frequencies, \( D(f) \propto 1/f \), so that there is no characteristic frequency (just as there is no characteristic length scale for spatial fractals).

1/f-Noise

Whenever the power spectral density, \( S(f) \), scales as \( 1/f \), the system is said to exhibit 1/f noise (or flicker-noise). Despite being found almost everywhere in nature – 1/f-noise has been observed in the current fluctuations in a resistor, in highway traffic patterns, in the price fluctuations on the stock exchange, in fluctuations in the water level of rivers, to name just a few instances – there is currently no fundamental theory that adequately explains why this same kind of noise should appear in so many diverse kinds of systems. What is clear is that since the underlying dynamical processes of these systems are so different, the common bond cannot be dynamical in nature, but can only be a kind of "logical dynamics" describing how a system's degrees-of-freedom all interact. SOC may be a fundamental link between temporal scale invariant phenomena and phenomena exhibiting a spatial scale invariance. Bak, et. al., argue that 1/f noise is actually not noise at all, but is instead a manifestation of the intrinsic dynamics of self-organized critical systems.

A Possible Connection with Land Combat?

A simple way to test for SOC is to look for the appearance of any characteristic power-law distributions in a system's properties.
Richardson [266] and Dockery and Woodcock [77] have both reported linear SOC-like power-law scaling in land combat. Richardson examined the relationship between the frequency of "deadly quarrels" versus fatalities per deadly quarrel using data from wars ranging from 1820 to 1945. Dockery and Woodcock used casualty data for military operations on the western front after Normandy in World War II and found that the log of the number of battles with casualties greater than a given number C also scales linearly with log(C); see figure 32.

Figure 32. Analysis of WWII casualty data on the western front after Normandy (Dockery and Woodcock, [77])

The paucity of historical data, however, coupled with the still controversial notions of SOC itself, makes it difficult to say whether these suggestive findings are indeed pointing to something deep that underlies all combat or are merely "interesting" but capture little real substance. Even if the results quoted above do capture something fundamental, they apply only to a set of many battles. The problems of determining whether, or to what extent, a power-law scaling applies to an individual battle or to a small series of battles, and – perhaps most importantly – what tactically useful information can be derived from the fact that power-law scaling exists at all, remain open.

Complex Adaptive Systems

In simplest terms, complex adaptive systems (CASs) are complex systems (meaning that they consist of many nonlinearly interacting parts) whose parts can adapt to changing environments. Moreover, each "part" typically exists within a nested hierarchy of parts within parts; see figure 33.
Traditionally, simulations of complex systems have consisted of mathematical or stochastic models, typically involving differential equations, that relate one set of global parameters to another set and describe the system's overall dynamics. The behavior of a system is then "understood" by looking at the relationship between the input and output variables of the simulation. While such an approach is adequate for systems with parts that possess little or no internal structure, it is largely incapable of describing groups, or societies, in which the internal dynamics of the constituent members of the system represent a vital part of the underlying dynamics.

Additional drawbacks of traditional simulation methods include:

- a failure to distinguish among different levels of activity within real complex systems; that is to say, a failure to appreciate that global parameters, such as the population size of an ecology, are often profoundly related to local parameters, such as the decision-making processes of individuals within the ecology – traditional simulation methods, particularly those relying on a differential equation approach, seldom take into account this local-global dichotomy;

- an inability to analytically account (such as in a differential equation form) for individual actions and /or strategies of the constituent elements of a complex system;

- an inability to realistically account for the qualitative information that individuals may use in formulating their strategies and upon which they may base their local decisions

An alternative agent-based approach, described below, is to respect the nested hierarchy of dynamics and dynamical "decisions" that are made
in these complex systems, and to include a model of the
decision-making ability and adaptability of the constituent agents.

This section provides a brief introduction to complex adaptive systems
and agent-based simulations. The recent monograph on complex
adaptive systems by Holland [144] is an excellent overall source of
reference. Additional source material can be found in conference
proceedings edited by Hillebrand and Stender [137], Meyer and
Wilson [213], and Varela [307], and in a collection of papers edited by
Maes [202] (the latter reference provides both theory and practical
descriptions of the design of autonomous agents). A recent overview of
adaptive autonomous agents, including a discussion of open problems,
is also given by Maes [204].

**Characteristics**

Most complex adaptive systems share seven basic characteristics (see
Holland, [144]):

- Four properties:

  - *aggregation*
    - type I – reduction of pertinent variables by aggregating
      "similar" things into categories; identifying details that
      are unimportant for the problem at hand, categories
      consist of things that differ only in those irrelevant
      properties
    - type II – emergence of complex large-scale behaviors
      from the aggregate interactions of less complex agents
      (example: Hofstadter’s "Ant Fugue"); agents →
      meta-agents → meta-meta-agents → etc.

  - *nonlinearity* – if \( f \) is a function or an operator, and \( x \) is a
    system input (either a function or variable), then the effect
    of adding two inputs, \( x_1 \) and \( x_2 \) first and then operating on
    their sum is, in general, not equivalent to operating on two
    inputs separately and then adding the outputs together;

  - *information flows* – defined by nodes, connections and
    resources, any of which can change over time

  - *diversity* – diversity is neither accidental nor random; the
    persistence of any individual agent depends on the context
    provided by the other agents; roughly, each agent fills a
    niche defined by the interactions centering on that agent; if
    you remove one agent from the system – creating a "hole" –
the system typically responds with a cascade of adaptations resulting in a new agent that fills that hole

- Three mechanisms:
  - *tagging* – a mechanism that facilitates aggregation; used to manipulate symmetries – enabling CASs to ignore certain details while directing their attention to certain others; tags allow agents to select among other agents or objects that would otherwise be indistinguishable
  - *internal models* – used by agents to "anticipate" and "predict" events in their environment
  - *building blocks* – primitives used in building internal models

**Agent-Based Simulations**

Agent-based simulations of complex adaptive systems are predicated on the idea that the global behavior of a complex system derives entirely from the low-level interactions among its constituent agents. By relating an individual constituent of a complex adaptive system to an agent, one can simulate a real system by an artificial world populated by interacting processes. Agent-based simulations are particularly adept at representing real-world systems composed of individuals that have a large space of complex decisions and/or behaviors to choose from.

Lessons about the real-world system that an agent-based simulation is designed to model can be learned by looking at the emergent structures induced by the interaction processes taking place within the simulation.

The purpose behind building such simulations is twofold: it is to learn both the *quantitative* and *qualitative* properties of the real system. Agent-based simulations are well suited for testing hypotheses about the origin of observed emergent properties in a system. This can be done simply by experimenting with sets of initial conditions at the micro-level necessary to yield a set of desired behaviors at the macro-level. On the other hand, they also provide a powerful framework within which to integrate ostensibly "disjointed" theories from various related disciplines. For example, while basic agent-agent interactions may be described by simple physics and sociology, the internal decision-making capability of a single agent may be derived, in part, from an understanding of cognitive psychology.


Adaptive Autonomous Agents

The fundamental building block of most models of complex adaptive systems is the so-called adaptive autonomous agent. Adaptive autonomous agents are parts of a complex adaptive system that try to satisfy a set of goals (which may be either fixed or time-dependent) in an unpredictable and changing environment. They are "adaptive" in the sense that they can use their experience to continually improve their ability to deal with shifting goals and motivations. They are "autonomous" in that they operate completely autonomously, and do not need to obey instructions issued by a God-like oracle.

Depending on the system being modeled and the environment that an agent populates, an adaptive autonomous agent can take on many different forms. In Deneubourg et al.'s [213] study of decentralized collective sorting, for example, which was used earlier as an example of emergence (see page 91), the agents of the system are simple (nonadaptive) robots that move about their physical environment and make elementary decisions about whether to pick up or drop an object. Examples of adaptive agents populating "cyberspace" are the so-called "software agents" (or "knobots") that are entities that navigate computer networks or cruise the World-Wide-Web searching for relevant bits of data.

In general, an adaptive autonomous agent is characterized by the following properties:

- it is an entity that, by sensing and acting upon its environment, tries to fulfill a set of goals in a complex, dynamic environment

- it can sense the environment through its sensors and act on the environment through its actuators

- it has an internal information processing and decision making capability

- its anticipation of future states and possibilities, based on internal models (which are often incomplete and/or incorrect), often significantly alters the aggregate behavior of the system of which an agent is part

- an agent's goals can take on diverse forms:
  - desired local states
  - desired end goals
selective rewards to be maximized

internal needs (or motivations) that need to be kept within desired bounds

The adaptive mechanism of an adaptive autonomous agent is typically based on a genetic algorithm (see page 93).

**What Distinguishes the Study of Adaptive Autonomous Agents from Traditional AI?**

At first sight, the kinds of problems best suited for agent-based simulations appear to be similar to the kinds of problems for which traditional artificial intelligence (AI) techniques have been developed. *How is an agent-based simulation different from a traditional artificial intelligence approach?* Maes [204] lists these key points that distinguish traditional AI from the study of adaptive autonomous agents:

1. traditional AI focuses on systems exhibiting isolated "high-level" competencies, such as medical diagnoses, chess playing, and so on; in contrast, agent-based system target lower-level competencies, with high-level competencies emerging naturally, and collectively, of their own accord

2. traditional AI has focused on "close systems" in which the interaction between the problem domain and the external environment is kept to a minimum; in contrast, agent-based systems are "open systems," and agents are directly coupled with their environment

3. most traditional AI systems deal with problems in a piecemeal fashion, one at a time; in contrast, the individual agents in an agent-based system must deal with many conflicting goals simultaneously

4. traditional AI focuses on "knowledge structures" that model aspects of their domain of expertise; in contrast, an agent-based system is more concerned with dynamic "behavior producing" modules. It is less important for an agent to be able to address a specific question within its problem domain (as it is for traditional AI systems) than it is to be flexible enough to adapt to shifting domains
What Distinguishes Traditional Modeling Approaches from Agent-Based Simulations?

Fundamentally, an agent-based approach to modeling complex systems differs from more traditional differential-equation based approaches in that it represents a shift from force-on-force attrition calculations to considering how high-level properties and behaviors of a systems emerge out of low level rules. The conceptual focus of agent-based models shifts from finding a mathematical description of an entire system to a low-level rule-based specification of the behavior of the individual agents making up that system.

Table 6 compares the traditional reductionist approach to modeling complex systems with complex adaptive system/agent-based simulations.

Table 6. Comparison between traditional and agent-based approaches to complex systems modeling

<table>
<thead>
<tr>
<th>Traditional (Reductionist) Approach</th>
<th>Agent-Based Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>degrees-of-freedom</strong></td>
<td>relatively few</td>
</tr>
<tr>
<td><strong>interactions</strong></td>
<td>typically weak and linear; need to be hard-wired into model</td>
</tr>
<tr>
<td><strong>characteristic length and time scales</strong></td>
<td>$\approx 1$</td>
</tr>
<tr>
<td><strong>specification of complex boundary conditions</strong></td>
<td>can be difficult to specify analytically (say, as part of a partial differential equation model)</td>
</tr>
<tr>
<td><strong>model of individual combatant?</strong></td>
<td>necessarily crude; assumes that all combatants are the same</td>
</tr>
<tr>
<td><strong>aggregation of variables</strong></td>
<td>simpleminded aggregation of low-level variables</td>
</tr>
<tr>
<td><strong>long term behavior</strong></td>
<td>solve for steady-state equilibrium solution</td>
</tr>
<tr>
<td><strong>sought-for behavior</strong></td>
<td>is either accounted for explicitly or is typically absent; focuses on force-on-force attrition ratios</td>
</tr>
</tbody>
</table>

103
**Swarm**

Swarm is a multi-agent simulation platform for the study of complex adaptive systems. It is currently under development at the Santa Fe Institute.\(^\text{14}\)

The goal of the Swarm project is to provide the complex systems theory research community with a fully general-purpose artificial-life simulator. The system comes with a variety of generic artificial worlds populated with generic agents, a large library of design and analysis tools and a "kernel" to drive the actual simulation. These artificial worlds can vary widely, from simple 2D worlds in which elementary agents move back and forth to complex multi-dimensional "graphs" representing multidimensional telecommunication networks in which agents can trade messages and commodities, to models of real-world ecologies.

Swarm has been intentionally designed to include as few ad-hoc assumptions about the design of a complex system as possible, so as to provide a convenient, reliable and standardized set of software tools that can be tailored by researchers to specific systems.

Though the prototype has been written using the C programming language, it is object-oriented in style. Future versions of Swarm will be implemented using the Objective-C language. Objective-C is an object-oriented extension of the C language that is widely available as part of the GNU C compiler, and is available on the World-Wide-Web.

Everything in Swarm is an object with three main characteristics: Name, Data and Rules. An object's Name consists of an ID that is used to send messages to the object, a type and a module name. An object's Data consists of whatever local data (i.e. internal state variables) the user wants an agent to possess. The Rules are functions to handle any messages that are sent to the object. The basic unit of Swarm is a "swarm": a collection of objects with a schedule of event over those objects. Swarm also supplies the user with an interface and analysis tools.

The most important objects in Swarm, from the standpoint of the user, are agents, which are objects that are written by the user. Agents represent the individual entities making up the model; they may be ants, plants, stock brokers, or combatants on a battlefield. Actions consist of a message to send, an agent or a collection of agents to send the message, and a time to send that message. Upon receiving a  

\(^{14}\) This section is based on the papers "An Overview of the Swarm simulation system," by '94 Swarm Team, Santa Fe Institute and "The SWARM simulation system and individual-based modeling," by D. Hiebler.

\(^{15}\) World-Wide-Web URL link = http://www.santafe.edu/projects/swarm/.
message, agents are free to do whatever they wish in response to the message. A typical response will consist of the execution of whatever code the user has written to capture the low-level behavior of the system he is interested in. Agents can also insert other actions into the schedule.

Three other properties of Swarm are noteworthy:

1. *Hierarchy.* In order to be better able to simulate the hierarchical nature of many real-world complex systems, in which agent behavior can itself be best described as being the result of the collective behavior of some swarm of constituent agents, Swarm is designed so that agents themselves can be swarms of other agents. Moreover, Swarm is designed around a time hierarchy. Thus, Swarm is both a nested hierarchy of swarms and a nested hierarchy of schedules.

2. *Parallelism.* Swarm has been designed to run efficiently on parallel machine architectures. While messages within one swarm schedule execute sequentially, different swarms can execute their schedules in parallel.

3. *Internal Agent Models.* One can argue that agents in a real complex adaptive system (such as the economy) behave and adapt according to some internal model they have constructed for themselves of what they believe their environment is really like. Sometimes, if the environment is simple, such models are fixed and simple; sometimes, if the environment is complex, agents need to actively construct hypothetical models and testing them against a wide variety of assumptions about initial states and rules and so forth. Swarm allows the user to use nested swarms to allow agents to essentially create and manage entire swarm structures which are themselves simulations of the world in which the agents live. Thus, agents can base their behavior on their simulated picture of the world.

The many kinds of problems that Swarm is well suited for include economic models (with economic agents interacting with each other through a market), the dynamics of social insects, traffic simulation, ecological modeling, simulation games such as SimCity and SimLife, and general studies of complex systems, cellular automata, and artificial life.

**Neural Networks**

One might facetiously ask, "How can a three year old baby put a CRAY X-MP supercomputer to shame?" The very serious answer is that "She can recognize uncle Seymour's face infinitely faster!" No matter how
powerful a computer one has, no matter how powerful an imaging system and image recognition software one is using, it is a fact that a young child will be infinitely better at recognizing certain patterns than the state-of-the-art hardware/software combination. Aside from the obvious fundamental question "Why?", an equally important question is whether or not a child's internal processing is something that can itself be mimicked or even directly simulated.

Part of the answer may lie in what the question itself tacitly assumes: how a child processes information is distinctly different from the way traditional computers process information. One is obviously parallel, the other is serial; one is algorithmic, diligently following a specific set of instructions one instruction at a time, the other is essentially "free form," organizing information and defining computational route seemingly on-the-fly; etc. The full answer depends on how well we are able to navigate our way on an emerging paradigm shift in the way computation is itself understood.

Neural nets (NNs) represent a radical new approach to computational problem solving. The methodology they represent can be contrasted with the traditional approach to artificial intelligence (AI). Whereas the origins of AI lay in applying conventional serial processing techniques to high-level cognitive processing like concept-formation, semantics, symbolic processing, etc. – or in a top-down approach – neural nets are designed to take the opposite – or bottom-up – approach. The idea is to have a human-like reasoning emerge on the macro-scale. The approach itself is inspired by such basic skills of the human brain as its ability to continue functioning with noisy and/or incomplete information, its robustness or fault tolerance, its adaptability to changing environments by learning, etc. Neural nets attempt to mimic and exploit the parallel processing capability of the human brain in order to deal with precisely the kinds of problems that the human brain itself is well adapted for.

There is a strong connection between cellular automata (see page 81) and neural networks. Fundamentally, CA represent a paradigm whereby the conventional emphasis of looking for the origins of complex behaviors in sets of "complex" building blocks is shifted to an entirely different mode-of-thought in which complexity itself is viewed as an emergent phenomenon built upon an assemblage of possibly very "simple" parts. From a purely philosophical point of view, it could also be argued that there is no better known example of a truly emergent phenomenon than that of the emergence of consciousness out of the large network of functionally "simple" (and certainly unconscious) neuronal components that make up the human brain. Now, while we have been inexcusably cavalier in our usage of terms like "complex" and "simple" – no respectable neurophysiologist would ever call a neuron simple! – it is safe to say that neural nets arguably represent the prototypical complex system.
Reduced to their essentials, CA are dynamical systems consisting of a discrete cellular state space, the nodes of which contain one of a finite number of discrete values, and evolve synchronously in time according to local update rules. Very crudely speaking, a biological neural network likewise consists of a large space of interconnected nodes whose dynamical behavior is a local function of other nodes to which it is connected. Artificial neural nets can be loosely thought of as being nothing more than a set of biologically inspired CA rules.

There exist many excellent books and collections of papers on this broad and rapidly growing subject. A recent book by Jubak [163] provides a good nontechnical introduction. The collections edited by Anderson and Rosenfeld [5] and Shaw and Palm [284] contain most of the important early landmark papers. Some of the better texts are those by Hecht-Nielson [130], Hertz, Krogh and Palmer [135], and Peretto [247]. There are also journals that specialize in neural nets such as Neural Computation, published by MIT and Neural Networks, published by Pergamon.

**A Short History**

Table 7 lists some developments in neural net research. This list is by no means exhaustive and is intended only to highlight some of the key events. There are four main points to be taken from this table:

<table>
<thead>
<tr>
<th>Year</th>
<th>Developers</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>1943</td>
<td>McCulloch-Pitts</td>
<td>first neuron model</td>
</tr>
<tr>
<td>1949</td>
<td>synaptic learning rule</td>
<td>Hebb</td>
</tr>
<tr>
<td>1958</td>
<td>simple perceptron model</td>
<td>Rosenblatt</td>
</tr>
<tr>
<td>1960</td>
<td>least-mean-square (LMS)/Delta-rule</td>
<td>Widrow and Hoff</td>
</tr>
<tr>
<td>1969</td>
<td>Perceptrons (a critical look at what neural nets can and cannot do)</td>
<td>Minsky and Papert</td>
</tr>
<tr>
<td>1982</td>
<td>autoassociation</td>
<td>Hopfield</td>
</tr>
<tr>
<td>1982</td>
<td>parallel distributed processing</td>
<td>Feldman, et. al.</td>
</tr>
<tr>
<td>1983</td>
<td>Boltzman machine</td>
<td>Hinton</td>
</tr>
<tr>
<td>1985</td>
<td>back-propagation learning rule</td>
<td>Rumelhart, et. al.</td>
</tr>
<tr>
<td>1990</td>
<td>i80170NX neural chip</td>
<td>Intel Corporation</td>
</tr>
</tbody>
</table>

1. The first serious work on neural nets dates back to 1943, so that while neural nets have been getting increasingly more press in recent years, it cannot be said that they are of "recent origin."
2. There is a thirteen year gap between the publication of Minsky and Papert's enormously influential book called *Perceptrons* in 1969 and the publication of a mainstream neural net physics paper by Hopfield in 1982, emphasizing the idea of memories as dynamically stable attractor states. This is not necessarily meant to imply that neural net research had not been going on in the intervening years, but it is certainly true to say that Minsky and Papert's powerful critique discouraged more than a few researchers from entering the (at that time) fledgling field.

3. Long stretches of time sometimes pass between the introduction of a basic design and when that design gains a more practical utility with a concomitant learning rule. For example, while it was known fairly early on that some of the limitations of Rosenblatt's simple perceptron could be overcome by adding "hidden" neurons, it was not until Rumelhart,*et al.*'s backpropagation rule was introduced in 1985 that an appropriate learning rule was finally found.

4. Just as for the study of cellular automata in general, for which software is nice to have for prototyping and preliminary study but are clearly inadequate for any large-scale simulation (lattice gases, for example, would have remained nothing more than an interesting theoretical exercise had it not been for dedicated hardware simulators like RAPI and CAM-6), neural nets do not really come into their own as *problem solvers* unless their designs are hard wired into silicon. The Intel chip listed in table 7 is but one example of a growing number of increasingly more powerful and fully programmable chips appearing on the commercial market.

**An Heuristic Discussion**

Neural nets are designed to exploit the most powerful computational characteristics of the human brain, most notably its efficiency at pattern recognition. As mentioned earlier, for example, a three-year-old child is considerably more adept at recognizing objects and faces than even the most advanced artificial intelligence system running on a top-of-the-line supercomputer. Moreover, the brain is robust and fault-tolerant, easily deals with probabilistic, fuzzy or even inconsistent information, and rapidly adapts to a changing environment by "learning," and does so without having to be continuously "re-programmed." Although the analogy between brains and NNs is crude at best (for example, the nodes and interconnections between the nodes of a neural network are gross oversimplifications of the brain's actual structure), it is close enough that NNs are capable of closely mimicking some of the brain's own functionality. In this section we discuss a class of NNs known as *multi-layer, feed-forward NNs*
that use the backpropagation learning rule. Such NNs are particularly adept at solving general pattern recognition problems.

Consider the "simple" problem of correctly identifying a sequence of handwritten characters (i.e. digits and letters). From a conventional programming perspective, the task is more than a little bit daunting. Although each character typically has a large set of unequal (and possibly widely varying) representations, each of these representations is valid. Some characters may appear darker than others, for example; others may be hurriedly written so that some detail is washed out; still others may be written in script rather than block-printed. Unless the character-recognition software explicitly accounts for all possible variations (or all possible 'templates') of the representation of each character, it is doomed to be imperfect from its inception. Even if a massive look-up table is constructed, consisting of very many variations of each character, some decision module must nonetheless also be constructed to deal with samples that do not match any of the pre-computed entries of the look-up table. The character-recognition software can only be as good as its decision module. How is a character to be identified if the sample is smudged, for example? While we may take for granted the fact that our brain can easily ignore any "dirt" or "smudges" that appear on an input to correctly identify the underlying character, a smudged character can also easily be different enough such that the smudged character will remain effectively unrecognizable to the software.\(^{16}\)

Traditional software's basic problem — namely, how to reliably deal with noisy and/or imprecise input — is the typical neural-net's strength. While most software-based recognition schemes depend on pre-defined sets of contingency rules to provide specific responses to inputs that do not match any elements of the basic look-up table, neural-nets learn to generalize from a basic set of input data, and thus require no special programming to process noisy input. Just as traditional software can only be as good as the algorithm that it implements, a neural-net can only be as good as the data set on which it trains.

**Defining and Training a Neural Network**

A typical net consists of three layers: an input layer, an output layer and one or more hidden layers (see figure 34). The input layer is chosen to correspond in some way to the set of input data. For example, if the input is to consist of images of handwritten characters, one might use a

\(^{16}\) The difference between two patterns is typically measured by computing the Hamming Distance. The Hamming Distance between two binary strings \(S_1 = 001001110\) and \(S_2 = 010010111\) is defined to be the number of string-entries in which the digits are different. The Hamming distance between \(S_1\) and \(S_2\) is therefore 5. The Hamming distance between two handwritten characters, for example, may be defined as the number of on/off pixels by which their two respective black-and-white digitized images differ.
10 x 10 grid of neurons, with each neuron having either value 1 or 0 depending on whether or not a portion of the character overlays that neuron. Similarly, the output layer is chosen to correspond to the set of output data. If the output is to consist of the 26 letters of the alphabet and the 10 numerical digits, a natural choice would be to have a layer with 36 neurons corresponding to each possible character or digit output. Unfortunately, choosing the size and number of hidden layers involves a little bit of black-magic: aside from what are essentially heuristic aids, there are no formal theorems specifying an optimal design. It can be shown that (1) for the net to be able to solve nontrivial problems there should be at least one hidden layer, and (2) any net with more than two hidden layers is functionally equivalent to a net with two hidden layers. The total number of neurons within the hidden layers should (usually) be between 1/2 to 2 -1/2 times the number of input neurons. In practice, it is best to remember that a net with too few hidden neurons will be unable to learn what is required of it; a net with too many hidden neurons will tend to overgeneralize what it has learned.

Figure 34. Schematic representation of a multi-layer feed-forward neural network

Each of the nodes in the first hidden layer is connected to each input node; their values depend on the weights assigned to each of these connections. Likewise, each node of the second hidden layer (if one exists) is connected to each node of the first hidden layer, and so on until the last, or output, layer is reached (all of whose nodes are connected to each node of the last hidden layer). Learning is the process whereby the net adjusts the set of its internal weights so that for each input fact the output state corresponds to the desired output. The net is typically first run with a set of random weights so that its initial output bears little relation to the input. As it "re-looks" at the same set of input/output fact pairs many times, the net continuously readjusts the weights so as to bring its processed output closer and closer to the desired output. After some transient learning period that
depends on the size of the network and number of input facts, a final weight set is achieved such that the output to each fact is what it is required to be.

Using a Trained Net

Once learning is completed, and for the same set of facts with which it was trained, the net may be used as a simple data-retrieval program: for each fact in the training set, it will correctly reproduce the desired output. The net's real strength, however, lies in its ability to abstract and generalize from this training set and thereby deal with inputs that do not match any of the training facts. Having been trained on a particular set of 36 characters and digits, for example, corresponding to one trainer's unique handwriting, the net will then be able to correctly identify the characters input from a digitized sample of someone else's handwriting. Each input/output fact pair will have been generalized so that a much larger set of similar but unequal input handwriting samples can be recognized as corresponding to the same output.

While the choice for the input space in the character-recognition example may have been "obvious"—since we know that the output depends only on the input image, our only problem is to find some natural correspondence between a set of neurons and an arbitrary input image—another strength of NN technology is that a net will (using slightly anthropomorphic language) use whatever subset of input facts that it decides is really important for predicting the desired output. Even if the trainer is himself unsure of exactly what set of input facts are really important (we will outline an example in a moment), as long as he uses a list that is a super set of the list of facts that truly matter, the neural net will experience no particular difficulty with training. In other words, the net effectively learns to parse out and use only those facts that are relevant to reach the desired conclusions. Other facts, having nothing to do with the desired output, are acknowledged only by being given essentially zero weight. The net will recognize them to be unimportant and train itself to ignore them. By the same reasoning, a net can also suggest that certain facts that the trainer considers to be unimportant are in fact important in reaching a conclusion and should not be ignored.

Consider a neural net approach to predicting when the NN trainer himself will choose to take a coffee break. Obviously, many factors play a role, both on a conscious and unconscious level. The amount of time that has passed since the last break is clearly an important factor; but constraints such as how close his work is to a deadline or how much time is left until the analyst must leave for day are also important. Perhaps when the height of graph-paper on the left-hand-side of his desk exceeds some threshold, the analyst begins considering taking a break? On the other hand, the decision to take a break may have nothing at all to do with such factors as the day of the week, the
temperature in a neighboring office or the color of the analyst's socks. The input layer of a net designed to recognize the trainer-state/decision-to-break pattern may consist of as many neurons as analyst-state facts the trainer decides play an important role. One neuron, for example, could correspond to the time elapsed since the last break; another to how much time is left until the end of the day, etc. As long as the input space contains the set of facts that do play an important role in finally yielding either a yes-break or no-break decision from the analyst (a set which may indeed be a-priori unknowable), the net will learn to ignore the unimportant part of the input space. The output may consist of a single neuron, which takes the value 1 when the decision is to take a break, and value 0 when the decision is to keep working.

General Model Development: A Short Primer

While the details of designing a neural net solution to a particular problem can be quite involved, the basic strategy is fairly simple. The important point to remember is that designing a net is almost an antithesis of conventional programming. There are no rules or algorithms to write (except for the underlying code defining the learning algorithm, of course). Instead, the effort that is conventionally put into the programming end of a solution is replaced by the effort that must be put into constructing a sample solution set, one that must often be put together without an explicit or an a-priori knowledge of the method of solution. For example, using the example of image recognition ("Uncle Seymour's face"), while most of us are instantly able to recognize the faces of even the most casual of our acquaintances, very few of us are able to describe exactly how we are able to accomplish this task.

Fundamentally, all feed-forward nets follow the same basic steps of a model development cycle:

1. define the problem
2. define the input-output fact set
3. define the neural net structure
4. train
5. test

It cannot be stressed strongly enough that the first step, defining the problem, is far from being a simple task. Great care must be taken to identify precisely what one wishes for the net to "learn."
There is a telling story about how the army recently went about teaching a NN to identify tanks set against a variety of environmental backdrops. The programmers correctly fed their multi-layer net photograph after photograph of tanks in grasslands, of tanks in swamps, of tanks surrounded by trees, of hills without tanks, and so on. The idea was for the net to get a broad enough sampling of scenes with tanks both present and absent so as to be able to tell, in general, whether a tank was or was not present in an arbitrary image. After many trials and many thousands of iterations, the NN finally learned all of the images in the carefully prepared database. When the presumably trained net was tested on other images that were not part of the original training set, it failed to do any better than what would be expected by chance alone. The problem was that the input/training fact set was inadvertently statistically corrupt. The database consisted mostly of images that showed a tank only if there were heavy clouds, or the tank itself was immersed in shadow or there was no sun at all. The Army’s neural net had indeed identified a latent pattern, but it unfortunately had nothing to do with tanks: it had effectively learned to distinguish bright from not-so-bright scenes.

The obvious lesson to be taken away from this amusing example is that how well a net “learns” the desired associations depends almost entirely on how well the database of facts is defined.

Once the input-output fact pair has been put together, the next challenge is to find an appropriate net design; i.e. to determine how many input and output neurons should be used and how many hidden layers should be placed between them. Typically, the form of the input-output facts in the database determine the number (and type) of input and output neurons. In the handwriting recognition example we used earlier in which a net is to learn the 26 letters of the alphabet, for example, a natural choice was to use $N \times M$ input neurons to encode an input image and 26 output neurons, each of which corresponds to a given letter. If, instead, the problem is to construct a financial predictor, where the input data consists of such facts as the consumer price index, the price of crude oil, the unemployment rate, and so on, and the desired output is an estimate of the Dow-Jones stock average, it is natural to design a net that has as many input neurons as there are available input facts and one output neuron whose value is equal to the predicted Dow-Jones average. Countless other examples could of course also be imagined. The point is that once the problem has been carefully defined and the available information structured in some form, the number of input and output neurons is essentially determined.

**Backpropagation Algorithm**

The backpropagation learning rule (also called the generalized delta rule) is credited to Rumelhart and McClelland [271]; refer to figure 34.
for a schematic of a multi-layered neural net's structure. Notice that the design shown, and the only kind we will consider in this section, is strictly feed forward. That is to say, information always flows from the input layer to each hidden layer, in turn, and out into the output layer. There are no feedback loops anywhere in the system.

One or more hidden layers are sandwiched between the input and output layers and, for the moment, consist of an arbitrary number of neurons. While there are, unfortunately, no rigorous theorems specifying what number should be used for a given problem, useful heuristics do exist.

The backpropagation learning rule gives a prescription for adjusting the initially randomized set of synaptic weights (existing between all pairs of neurons in each successive layer) so as to minimize the difference between the neural net's output for each input fact and the output with which the given input is known (or desired) to be associated. The backpropagation rule takes its name from the way in which the calculated error at the output layer is propagated backwards from the output layer to the Nth hidden layer to the (N-1)th hidden layer, and so on. Because the learning process requires us to "know" the correct pairing of input-output facts beforehand, this type of weight adjustment is called supervised learning.

**Pseudo-Code**

Without derivation, we now present a seven-step pseudo-code implementation of the backpropagation learning rule. It is to be applied for each pattern ". Assume that we have a neural net with L layers (1 = 1, 2, ..., L). Let \( h_i^l \) represent the output of the \( i \)th neuron in the \( l \)th layer; \( h_i^0 \) is therefore equal to the \( i \)th input, \( \sigma_i \). The weight of the connection between \( h_j^{l-1} \) and \( h_i^l \) is labeled \( w_{ji}^l \).

- **Step 1:** Initialize all weights to small random values.

- **Step 2:** Set the input layer equal to the input values for the first input/output fact pair: i.e. let \( h_k^0 = \sigma_k \) for all values of \( k \).

- **Step 3:** Propagate the input signal forward through the various layers of the net; i.e. calculate \( h_i^l = f_\alpha(\sum_j w_{ij}^l h_j^{l-1}) \), where \( f_\alpha(x) = \frac{1}{1 + e^{-\alpha x}} \) is a sigmoidal threshold function and \( \alpha \) is a parameter added to control the steepness of the curve.

- **Step 4:** Calculate the differences \( \Delta_i^l \) for the output layer: \( \Delta_i^l = f'_\alpha(h_i^l)(\sum_j \delta_j^l - h_i^l) \), where \( f'_\alpha(x) \) is the derivative of \( f_\alpha(x) \), is the net's calculated output and \( \delta_i^l \) is the actual output of the net,
Step 5: Obtain the differences $\Delta$'s for each of the preceding layers by propagating the errors backwards:
$$
\Delta_{l-1} = f'(h_{l-1})\sum_j w_{ij}^l \Delta_j^l, \quad l = L, L-1, ..., 2.
$$

Step 6: Adjust the weights according to $w_{ij} \rightarrow w_{ij} + \delta w_{ij}$, where
$$
\delta w_{ij} = \eta \Delta_i^l h_{ij}^{l-1}
$$
and $\eta$ is an adjustable learning constant.

Step 7: Go back to step 2 and repeat for the next pattern. Stop when the difference between the computed and desired output is less than some pre-assigned threshold.

The basic backpropagation algorithms is in practice often very slow to converge. Moreover, it can sometimes get stuck in undesired spurious attractor states. This is an unfortunate artifact that plagues all cost-function minimization schemes. In recent years, however, a number of alternative formulations to improve convergence have been suggested. Consider the learning constant $\eta$, that effectively determines how fast the system moves down a "hill" of the energy surface. Although smaller values of $\eta$ lend stability, they also tend to slow down the convergence to unreasonably slow rates of convergence. On the other hand, if $\eta$ is too large, the algorithm tends to oscillate and become unstable. Among the methods suggested to alleviate these problems are (1) using successively smaller values of the learning constant, (2) continuously adapting the value of the learning constant to how well the convergence is doing, and (3) adding a so-called momentum-term. A fourth method is to add a bit of noise at each step. The idea in this last method is to use the noise to knock the system out of an undesired local minima.

Since their introduction, feed-forward backpropagating neural nets have been used to "solve" a wide range of interesting problems, striking in their diversity. Applications include playing backgammon, recognizing hand-written zip-codes, financial bond rating, visual pattern recognition, classification of seismic signals, sonar target recognition, and navigating a car, among many others.

Example: NETtalk

An important and influential application of a multi-layered backpropagating neural net is NETtalk, designed by Sejnowski and Rosenberg in 1987. NETtalk learns to convert English text into speech and displays many of the characteristics of learning normally ascribed to human learning, including a power-law form for its learning curve and an increase in its ability to generalize as the size of its training set increases. Moreover, NETtalk is robust and fault tolerant. Its performance degrades gracefully and not catastrophically if its set of synaptic weights is damaged. Once trained, NETtalk is also able to
relearn a given fact set much faster after some damage has been done to it than if it had to start from scratch with the original training.

Generating speech from written text is a profoundly difficult problem to solve using conventional programming techniques. Part of the difficulty is due to the fact that words are not always pronounced according to how they are spelled. Although we are all taught the "rules" when we are young, as our experience grows we learn that each rule has its fair share of exceptions. Even a simple sentence like "This is a sentence," shows that spelling can be a poor cue for pronunciation. The first two s's are pronounced differently, but they both appear at the end of words and are both preceded by an i. The third s is pronounced the same way as the first but instead appears at the beginning of a word. NETtalk deals with this problem by looking at groups of letters to provide context sensitivity.

NETtalk consists of 203 input neurons, one hidden layer composed of 80 neurons and 26 output neurons. The input consists of a string of 203 letters, with a 7-site long window that slides over the text to provide the necessary context sensitivity for the net to be able to learn to pronounce the middle letter. The output consists of 26 elementary speech sounds called phonemes. Phonemes are similar to the pronunciation guides found in standard dictionaries. Several sets of phonemes are available.

Sejnowski and Rosenberg used two different sets of words for training: (1) 1024 words taken from phonetic transcriptions of informal continuous speech by children, and (2) a subset of the 1000 most commonly used words selected from Miriam Webster's Pocket Dictionary.

NETtalk managed to learn the informal speech database well enough after only a few training cycles (with each cycle being one complete pass through the database) to utter intelligible speech. Ten passes were sufficient for it to utter fully understandable words, and 50 cycles proved enough for NETtalk to attain a 95% accuracy rate. Initially, NETtalk was able to learn only gross features such as the difference between consonants and vowels. Since it always responded with the same vowel whenever any vowel was input and with the same consonant whenever a consonant was input, in this early phase NETtalk sounded like a babbling child. Gradually, NETtalk learned to recognize the boundaries between words and was thus able to begin uttering pseudowords. As its learning was further enhanced, NETtalk's output steadily improved to the point of intelligibility.

Perhaps the most striking result is an early demonstration of fault tolerance. When Sejnowski and Rosenberg artificially "damaged" the net by adding some amount to random noise to the synaptic weights, they found that NETtalk's ability to "speak" degraded only gradually and
not suddenly, as might be expected of a conventional rule based systems from which a subset of rules was suddenly deleted.

Other Designs

This brief survey has only touched upon one of the more familiar neural net designs, chosen mainly for its pedagogical value. There are of course a large number of other important models, some of which have wide applicability, some of which are optimized for a particular kind of problem:

- **Adaptive resonance.** One obvious drawback to using a backpropagating neural net is the need to retrain the net every time a new problem is added to the training set database. While the net should, in principle, have no problem in learning a new fact, there is the possibility that the newly trained net will forget previously stored information. This is sometimes loosely referred to as the stability-plasticity problem. One neural net design that addresses this concern is called *adaptive resonance*, and is due to Carpenter and Grossberg [41]. A critical feature of adaptive resonance is its ability to switch between a learning state in which the net's internal parameters can be modified (plasticity) and a fixed state wherein previously stored data cannot be damaged (stability).

- **Supervised learning.** The backpropagation algorithm assumes that the output part of a desired input-output set of pairs is known a-priori. In practice, of course, one often does not know the output. It is certainly reasonable to expect a net to effectively tell the trainer what latent patterns and similarities exist within a clump of data. For this one needs an entirely different neural net design, one that is optimized for finding common features across a range of input patterns in an unsupervised fashion. A well known exemplar of this class is Kohonen's *self-organizing feature map* [178]. A more general approach to unsupervised learning is called *competitive learning*, described by Rumelhart [271]. There are still other schemes, such as Hecht-Nielsen's *counterpropagation* networks, that combine supervised and unsupervised learning in one net.

- **Tailoring a design to a specific problem.** Algorithms such as backpropagation can be viewed as general purpose designs. By modifying a net's size, topology, data set, or energy function, such nets can be applied to a wide variety of problems. However, there are problems for which either such a general-purpose design does not suffice or for which a better, optimized, design can be constructed. One such example is Fukushima's *cognitron* and larger-scaled *neocognitron* neural net designs [104], which are
specifically tailored to recognize handwritten characters. Both the cognitron and neocognitron can learn with or without a supervisor and represent an approach based on exploiting the anatomy and physiology of the mammalian visual system.

Lessons of Complex Systems Theory

The major lesson of complex systems theory is that complex behavior is usually an emergent self-organized phenomenon built upon the aggregate behavior of very many nonlinearly interacting "simple" components. It advocates, in essence, a *simplicity breeds complexity* approach to the study of complex systems.

The critical points to remember are...

- **Nonlinearity.** Without nonlinear interactions there can be no deterministic chaos in simple systems and no complex behavior in complex systems. Moreover, nonlinear systems appear to be much more pervasive than linear systems. By virtue of nonlinearity, the behavior of the "whole" is not just a simple aggregate of the constituent "parts."

- **Interconnectivity.** How the parts of a complex system are interconnected is just as important as what those parts are and what does parts do.

- **Context/Wholeness.** The effect that parts have on the remainder of the system – literally, how those parts are defined within the complex system – is determined by the context of the whole within which those parts exist. In referring to any part P of a complex system, one must also point to various other parts with which P interacts (or may interact in the future).

- **Process:** Simple dynamical systems are characterized by simple attractors – fixed points, limit cycles, quasiperiodic and chaotic (or strange) attractors. Although one can also try to characterize the behavior of complex systems with these attractor "labels," such a description would entirely miss the essence of what it means to be a complex system. A complex system embodies process, a ceaseless search for a better "solution" for an ill-defined, amorphous -ever receding "problem." In Zen-like fashion, you can say that the harder one tries to pin-down the behavior of a complex system with some static measure, the further one is from understanding what the complex system is really doing.
Adaptability. The essence of a complex adaptive system is that its constituent parts are not Newtonian "billiards" that react blindly (but in well-defined fashion) to the world around them, but are instead endowed with an ability to sense, learn from, and adapt to their environment as they and the environment both evolve in time. A related lesson is that individual solutions (or evolutionary timelines) are essentially non-reproducible; a given system may "solve" a given problem in many different ways.

Emergence. Perhaps the central concept of complex systems theory is that high-level behaviors emerge naturally out a brewing soup of low-level interactions. A flock of birds (or "Boids," see page 73) does not need a central direction to behave in an apparently orchestrated manner. Nowhere on the lattice rule-level in Conways Life CA game (see page 87) is there any hint of the particle-like glider that spontaneously emerges on a higher level, and then apparently obey a dynamics all its own. The lesson is that where there is an assemblage of very many nonlinearly interacting parts, there is a good possibility of emergent behaviors on higher levels than those defining the underlying interactions. Moreover, such emergent behavior can appear on multiple spatial and temporal levels.
Land Warfare and Complexity Theory: Preliminary Musings

In this section we outline a few preliminary musings on the applicability of nonlinear dynamics and complex systems theory to the understanding and/or representation of land warfare. An in-depth analysis and discussion of the ideas outlined below will be provided in Part II of this report.\(^7\)

The fundamental question that is addressed, at least indirectly, in this report, and more fully in the follow-on paper, is "What does complexity theory tell us about land warfare?"

This question really embodies three separate but interrelated issues (see figure 35):

1. Complexity theory
2. Land warfare
3. Modeling/Simulation

Figure 35. Interrelated issues of addressing land warfare as a complex system

\textit{Complexity theory} refers to any and all conjectures, hypotheses, theories, experiments, mathematical models, etc. having to do with the understanding of complex systems exhibiting a complicated (i.e. chaotic) behavior. In particular, complexity theory is assumed to include both nonlinear dynamics and complex systems theory, the

\(^7\) Land Warfare and Complexity, Part II: An Assessment of the Applicability of Nonlinear Dynamics and Complex Systems Theory to the Representation of Land Warfare is scheduled to be delivered to sponsor for review 1 July, 1996.

121
latter including a multitude of sub-disciplines such as artificial life, cellular automata, genetic programming, neural networks, etc.

*Land warfare* embodies all of the myriad problems and issues of land warfare, including combat attrition, command and control, coordination, intelligence, tactics and strategy, training, etc.

*Modeling/Simulation* is a generic label for the overarching context within which possible interconnections between the tools and methodologies of complexity theory as well as the issues and problems of land warfare can be fully explored.

In this last category, the most important question to ask is "What do you expect to get out of a particular model?" A model that is designed to explicitly mimic reality as closely as possible in order to predict the outcome of real battles is very different from a simulation designed merely to act as a synthetic combat environment within which combatants can obtain "realistic" training. Sandwiched in between these two extremes of modeling lies another class of pseudo-realistic models designed to provide insights only into selected key elements of the general pattern of behavior on the battlefield. While complexity theory may potentially offer interesting insights into all three levels of modeling and/or simulation, it is likely to provide its strongest support to these middle-level models. A more in-depth discussion of this very important point will be given in part II of this paper.

If one main theme runs consistently throughout all of the preceding discussion in this report, it is that complexity theory embodies an enormously large set of concepts, mathematical tools and methodologies. Consequently, even interpreting the basic question of what complexity theory tells us about land warfare is not at all an easy task. Instead of asking "What does complexity theory tell us about land warfare?" a better question is "How does idea I or methodology M, born of complexity theory, help us to understand problem P or issue S in land warfare?" Moreover, possible answers to this question should not be confined to finding applications on only the tactical level of combat but should include the operational and strategic levels of warfare as well.

It is easy to imagine the most seductive application of complexity theory to land warfare, namely some 10-or-20-year's down-the-road artificial-life-like simulation of battle, complete with impressively realistic 3D virtual reality graphics. It is more likely, however, that the real, albeit less immediately seductive, applications will lie in the conceptual trenches, out of the way and beneath the surface, altering a field commander's frame of reference for seeing what is really happening on the battlefield, establishing new criteria for collecting data, improving the way information is processed and communicated, providing real-time tactical and strategic decision aids, and providing
tools for extracting and understanding any subtle and/or otherwise "hidden" patterns of behavior on the battlefield.

It must also be remembered – and appreciated – that complexity theory is still very much in its infancy. The Santa Fe Institute in New Mexico, for example, which is widely recognized as being a leading research center for complex systems, was founded just a decade ago in 1984. Moreover, many of the analytical tools and models developed for the study of complex systems, such as genetic algorithms, genetic programming and agent-based simulations, have been either developed or refined as part of the artificial-life research effort that itself sprang up only in 1987. Consequently, it would be grossly unfair to complex systems theory to expect to find a mature set of tools and methodologies at such an early stage of this burgeoning field's development.

Table 8. A sampling of tools from nonlinear dynamics and complex systems theory

<table>
<thead>
<tr>
<th>Nonlinear Dynamics/Chaos</th>
<th>Complex Systems Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>power spectra</td>
<td>cellular automata</td>
</tr>
<tr>
<td>fractal dimensions</td>
<td>genetic algorithms</td>
</tr>
<tr>
<td>Kolmogorov-Sinai entropy</td>
<td>genetic programming</td>
</tr>
<tr>
<td>Lyapunov exponents</td>
<td>neural networks</td>
</tr>
<tr>
<td>attractor reconstruction</td>
<td>self-organized criticality</td>
</tr>
<tr>
<td>time-delayed embedding</td>
<td>agent-based simulations</td>
</tr>
<tr>
<td>chaotic control</td>
<td>SWARM</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
</tr>
</tbody>
</table>

It is also safe to say that, at its current stage of development, much of complex systems theory is concerned primarily with simulations and simulation-engines – such as the Santa Fe Institute's SWARM (see page 114) – that run them. So much so that it is not entirely incorrect to think of "complex systems theory" as being synonymous with "agent-based simulation."

Table 8 lists some of the analytical tools of nonlinear dynamics and complex systems theory that were discussed in earlier sections of this report.

**Framing the Problem**

Figure 36 summarizes the more salient points of the overall discussion and is meant to provide a basic framework for forging a connection between complexity theory and land warfare.
Figure 36 shows that there are four levels of applicability of complexity theory: level-1, consisting of specific analytical and mathematical tools such as cellular automata, genetic algorithms, genetic programming, and so on; level-2, consisting of general simulation systems such as SWARM, within which complex systems can be modeled; level-3, consisting of observations of behavior of specific systems; and level-4, consisting of sets of universal behaviors, such as the principle of self-organized criticality (see page 101).

Ideally, of course, one would like to take whatever insights complexity theory has come up with, or will come with, on the highest level (level-4) and apply them directly to the issues and problems of land warfare. The fact that this is exceedingly unlikely to happen in the foreseeable future is due in no small measure to the fact that, as of this writing, there are precious few "universal behaviors" populating level-4. Indeed, as alluded to in an earlier section (see page 101), self-organized criticality is arguably the only existing holistic mathematical theory of self-organization in complex systems! Therefore, if there is anything at all that falls under the rubric of complexity theory that is generally applicable to the problems of land warfare, it will most likely consist of specific sets of tools applied to specific problems, along with whatever insights can be gained by using general-purpose simulators such as SWARM to act as simulation "engines." There remains the possibility that complexity theory might shed some light on how battlefields may be configured (or compelled to self-organize) to achieve a maximum adaptability to a changing environment.

Figure 36 also shows that there are four levels of land warfare to which the tools and methodologies of complexity theory can be applied: (1)
tactical, (2) operational, (3) strategic and (4) general strategic, which refers to the socio-political strategies that are followed over long periods of time and which can therefore span over several conflicts.

Finally, figure 36 illustrates that there are three levels on which complexity theory can be applied to land warfare:

- **Level-1.** This is the most basic metaphor level to which most general discussions have been heretofore confined. This level consists of constructing and elaborating upon similar sounding words and images that most strongly suggest a "philosophical resonance" between behaviors of complex systems and certain aspects of what happens on a battlefield. The Clauswitzian images of "fog of war" and "friction" come to mind immediately [24]. There is nothing wrong, *per se*, with confining a discussion to this level, but one must always be mindful of the fact that metaphors are easily abused and "philosophical resonances" do not imply real connections.

- **Level-2.** This is the pragmatic and/or experimental level on which real-world data is mined to confirm or deny that there is more to a possible connection between complexity theory and land warfare than mere "philosophical resonance" alone. The best work along these lines has so far been conducted by Tagarev, *et al.* [298] and is discussed briefly below. Tagarev, *et al.* provide evidence of deterministic chaos in tactical, operational and strategic dynamics of a wide class of military behavior.

- **Level-3.** This is the "workhorse" level on which specific methodology borrowed from complexity theory is applied directly to specific issues and problems of land warfare. This might not be as intellectually provocative or satisfying as making a direct, one-to-one mapping between universal patterns of behavior of complex systems in general and patterns of combat on the battlefield (although this is remotely conceivable in some form); however, using genetic algorithms to evolve tactics in real-time in the heat-of-battle is impressive nonetheless. Most of the ideas and conjectures outlined in the following sections fall squarely into this third level of connections.

### Chaos in combat models

A fundamental lesson of nonlinear dynamics theory is that one can almost always expect to find some manifestation of chaos whenever nonlinearities are present in the underlying dynamics of a model. This
fundamental lesson has potentially significant implications for even the simplest combat models.

Miller and Sulcoski [217], for example, report fractal-like properties and a sensitivity-to-initial conditions in the behavior of a discretized model of the Lanchester equations (augmented by nonlinear auxiliary conditions such as reinforcement and withdrawal/surrender thresholds).

Non-Monotonicities and Chaos

A 1991 RAND study [73] uncovered chaotic behavior in a certain class of very simple combat models in which reinforcement decisions are based on the state of the battle. The study looked at non-monotonic and chaos in combat models, where "monotonic behavior" is taken to mean a behavior in which adding more capabilities to only one side leads to at least as favorable an outcome for that side.

The presence of nonmonotonicities has usually been interpreted to mean that there is something wrong in the model that needs to be "fixed" and has been either treated as an anomaly or simply ignored. The main thrust of the RAND report is that, while non-monotonicities often do arise from questionable programming skills, there is a source of considerably more problematic non-monotonicities that has its origins in deterministic chaos.

The RAND study found that "a combat model with a single decision based on the state of the battle, no matter how precisely computed, can produce non-monotonic behavior in the outcomes of the model and chaotic behavior in its underlying dynamics." [73]

The authors of the report draw four basic lessons from their study:

- models may not be predictive, but are useful for understanding changes of outcomes based on incremental adjustments to control parameters

- scripting the addition of battlefield reinforcement (i.e. basing their input on time only, and not on the state of the battle) generally eliminates chaotic behavior

- one can identify input parameters figuring most importantly in behavior of non-monotonicities - these are the size of reinforcement blocks and the total number of reinforcements available to each side

- Lyapunov exponents are useful to evaluate a model's sensitivity to perturbations

126
In general, the RAND report [73] concludes that

"In any combat model that depends for its usefulness on monotonic behavior in its outcomes, modeling combat decisions based on the state of the battle must be done very carefully. Such modeled decisions can lead to monotonic behavior and chaotic behavior and the only sure ways (to date) to deal with that behavior are either to remove state dependence of the modeled decisions or to validate that the model is monotonic in the region of interest."

**Minimalist Modeling**

Dockery and Woodcock, in their massive treatise *The Military Landscape* [77], provide a detailed discussion of many different "minimalist models" from the point of view of catastrophe theory and nonlinear dynamics. Minimalist modeling refers to "the simplest possible description using the most powerful mathematics available and then" adds layers "of complexity as required, permitting structure to emerge from the dynamics." Among many other findings, Dockery and Woodcock report that chaos appears in the solutions to the Lanchester equations when modified by reinforcement. They also discuss how many of the tools of nonlinear dynamics (see table 8) can be used to describe combat.

Using generalized predator-prey population models to model interactions between military and insurgent forces, Dockery and Woodcock illustrate (1) the set of conditions that lead to a periodic oscillation of insurgent force sizes, (2) the effects of a limited pool of individuals available for recruitment, (3) various conditions leading to steady state, stable periodic oscillations and chaotic force-size fluctuations, and (4) the sensitivity to small changes in rates of recruitment, disaffection and combat attrition of simulated force strengths.

This kind of analysis can sometimes lead to counter-intuitive implications for the tactical control of insurgents. In one instance, for example, Dockery and Woodcock point out that cyclic oscillations in the relative strengths of national and insurgent forces result in recurring periods of time during which the government forces are weak and the insurgents are at their peak strength. If the government decides to add too many resources to strengthen its forces, the chaotic model suggests that the cyclic behavior will tend to become unstable (because of the possibility that disaffected combatants will join the insurgent camp) and thus weaken the government position. The model instead suggests that the best strategy for the government to follow is to use a moderately low level of military force to contain the insurgents at their peak strength, and attempt to destroy the insurgents
only when the insurgents are at their weakest force strength level of the cycle.¹⁸

Generalizations of Lanchester's equations

In 1914, Lanchester introduced a set of coupled ordinary differential equations as models of attrition in modern warfare. The basic idea behind these equations is that the loss rate of forces on one side of a battle is proportional to the number of forces on the other. In one form of the equations, known as the directed-fire (or square-law) model, the lanchester equations are given by the linear equations dR(t)/dt = -αₐ B(t) and dB(t)/dt = -αₖ R(t), where R(t) and B(t) represent the numerical strengths of the red and blue forces at time t, and αₐ and αₖ represent the constant effective firing rates at which one unit of strength on one side causes attrition of the other side's forces. An encyclopedic discussion of the many different forms of the lanchester equations is given by Taylor ([299], [300]).

While the lanchester equations are particularly relevant for the kind of static trench warfare and artillery duels that characterized most of World War I, they are too simple and lack the spatial degrees-of-freedom needed to realistically model modern combat. The fundamental problem is that they idealize combat much in the same way as Newton's laws idealize the real chaos and complexity ridden physics of the world. Likewise, almost all lanchester equation based attrition models of combat suffer from many basic shortcomings:

- determinism, whereby the outcome of a battle is determined solely as a function of the initial conditions, without regard for Clausewitz's "fog of war" and "friction"

- use of effectiveness coefficients that are constant over time

- static forces

- homogeneous forces with no spatial variation

- no combat termination conditions

- the assumption that target acquisition is independent of force levels

- no consideration of the suppression effects of weapons

- and so on ...
Perhaps the most important shortcoming of virtually all lanchester equation based models is that such models rarely, if ever, take into account the human factor; i.e. the psychological and/or decision-making capability of the individual combatant.

Generalizations of the lanchester equations have included:

- an analytical extension of Lanchester's equations to allow feedback between movement and attrition; this is discussed by Epstein [90]

- a general exploitation of the analogy between the form of the lanchester equations and Lotka-Voltera equations describing predator-prey interactions in natural ecologies

- Partial differential equations to include maneuver; primarily work done by Protopopescu at the Oak Ridge National Laboratory

- Fuzzy differential equations to allow for imprecise information; see Dockery, [74]

- Stochastic differential equations to describe attrition processes under uncertainty

One can speculate that there might be a way to generalize the lanchester equations to include some kind of an internal aesthetic. That is to say, to generalize the description of the individual combatants to include an internal structure and mechanism with which they can adaptively respond to an external environment. A discussion of this idea will be given in Part II of this report. See, for example, N. Smith's "Calculus of ethics," [289], [290].

**Demonstration of chaos in war using historical data**

An earlier section, describing work by Richardson [266] and Dockery and Woodcock [77], has already alluded to the possibility that certain gross combat attrition statistics appear to follow a power-law scaling very reminiscent of the characteristic fractal-like behavior observed in self-organized critical (SOC) systems (see page 101).

Tagarev, et. al. [298] also provides extensive historical evidence of chaos in tactical, operational and strategic levels of military activity. Tagarev, et. al. examine (1) US fixed-wing aircraft losses during the Vietnam war, (2) US Army casualties in western Europe during World War II, and (3) historical trends in US defense spending.
Nonlinear dynamics and chaos in arms-race models

G. Mayer-Kress ([96], [117] and [210]) has written many papers on nonlinear dynamics and chaos in arms-race models and has suggested approaches to socio-political issues. His approach is to analyze computational models of international security problems using nonlinear, stochastic dynamical systems with both discrete and continuous time evolution. Many of Mayer-Kress' arms-race models are based on models of population dynamics first introduced by L. F. Richardson after World War I [266].

Mayer-Kress finds that, for certain ranges of values of control parameters, some of these models exhibit deterministic chaos. In one generalization of a discrete version of Richardson's equations that models the competition among three nations, for example, Mayer-Kress finds that the two weaker nations will form an alliance against the stronger nation until the balance of power shifts [210]. The alliance formation factor and economical constraints induce nonlinearities into the model that result in multiple stable solutions, bifurcations between fixed point solutions and time-dependent attractors. He has also identified parameter domains for which the attractors are chaotic.

Combat simulation using cellular automata

If one abstracts the essentials of what happens on a battlefield, ignoring the myriad layers of detail that are, of course, required for a complete description, one sees that much of the activity appears to involve the same kind of simple nearest-neighbor interactions that define cellular automata (See page 81). Woodcock, Cobb and Dockery [323] in fact show that highly elaborate patterns of military force-like behavior can be generated with a small set of cellular automaton-like rules.

In Woodcock, et. al.'s model, each combatant – or automaton – is endowed with a set of rules with which it can perform certain tasks. Rules are of four basic varieties:

- **Situation Assessment**, such as the determination of whether a given automaton is surrounded by friendly or enemy forces

- **Movement**, to define when and how a given automaton can move; certain kinds of movement can only be initiated by threshold and/or constraint criteria

- **Combat**, which governs the nature of the interaction between opposing force automata; a typical rule might be for one
automaton to "aim fire" at another automaton located within some specified fight radius

- **Hierarchical Control**, in which a three-level command hierarchy is established; each lower-level echelon element keys on those in the next higher echelon on each time step of the evolution.

These basic rules can then be augmented by additional rules to (1) simulate the impact that terrain barriers such as rivers and mountains have on the movement of military forces; (2) provide a capability for forces to respond to changing combat conditions (for example, a reallocation of firepower among three types of weapons: aimed firepower, area firepower and smart weapons firepower), and (3) replace entities lost through combat attrition. Figure 37 shows a schematic of three sample rules. A further extension involves relating notional features of battlefield geometry to the structure of real battlefields [77].

**Figure 37. Three sample rules in Woodcock, et. al.'s CA combat model**

Woodcock, et. al. stress that the goal of CA-based model of combat is *not* to codify a body of rules that comes as close as possible to the actual behavioral rules obeyed by real combatants; rather, the goal lies in "finding the simplest body of rules that both can generate nontrivial global combat-like phenomena and provide a new understanding of the combat process itself by extracting the maximum amount of behavioral complexity from the least complicated set of rules." [323] Additional details are discussed in chapters 3.1 and 3.2 of reference [77].

**Computer viruses ("computer counter-measures")**

A computer virus can be thought of as an autonomous agent. It is a computer program that tries to fulfill a goal or set of goals without the
intervention of a human operator. Typically, of course, viruses have rather simple and sinister goals of tampering with the normal operation of a computer system and/or computer network and then reproducing in order to spread copies of themselves to other computers. Computer viruses are particularly interesting to artificial life researchers because they share many of the properties of biological viruses.

From a military standpoint, computer viruses can be used in two ways: (1) as computer countermeasure agents to infiltrate enemy systems, or (2) as constructive "cyberspace allies" that, for example, can be programmed to maintain the integrity of large databases.

**Intelligent Software Agents**

Anyone who has spent even a small amount of time "surfing" the World-Wide-Web for information can attest to how difficult it is to find useful information. To be sure, the WWW is filled with untold numbers of glossy pages overflowing with all kinds of information. A quick use of a web search-engine such as *Lycos* usually suffices to uncover some useful sites. But what happens when one needs to find some information about a particularly obscure subject area? And what happens when one begins relying on one's web connection for more and more of one's daily workload: e-mail, stock quotes, work scheduling, selection of books, movies, travel arrangements, video conferencing, and so on?

A powerful emerging idea that helps the human "web-surfer" deal with this increasing workload and that is based in part on the methodologies of autonomous agents and genetic algorithms, is that of *Intelligent Software Agents*. Software agents are programs that essentially act as sophisticated personal assistants. They act as intermediaries between the interests of the user and the global information pool with which the user has traditionally dealt directly. Software agents engage the user in a cooperative process whereby the human operator inputs interests and preferences and the agent monitors events, performs tasks, and collects and collates useful information. Because software agents come endowed with an adaptive "intelligence," they become gradually more effective at their tasks as they begin learning the interests, habits and preferences of the user.

From a military standpoint, intelligent software agents can be used for *adaptive information filtering and integration* and as *tactical picture agents*, scouring and ordering the amorphous flood of battlefield and intelligence data.

---

19 See, for example, the collection of articles in *Communications of the ACM*, Volume 37, No. 7, July 1994.
Agent-based simulations

For many obvious reasons, the most natural application of complexity theory to land warfare is to provide an agent-based simulation of combat. The basic idea is to model land combat as a co-evolving ecology of local-rule-based autonomous adaptive agents.

An Irreducible Semi-Autonomous Adaptive Combat Agent (ISAACA) represents a primitive combat unit (infantryman, tank, transport vehicle, etc.) that is equipped with the following characteristics (see figure 38):

- a default local rule set specifying how to act in a generic environment (i.e. an embedded "doctrine")
- goals directing behavior ("mission")
- sensors generating an internal map of environment ("situational awareness")
- an internal adaptive mechanism to alter behavior and/or rules; adaptation is genetic-algorithm-based (see page 99) – each ISAACA effectively plays out a scenario using a genetically-encoded set of possible tactics; where fitness is the expected payoff with respect to some internal value system

An ISAACA collective, represented schematically in figure 39, consists of local and global commanders, each with their own command radii, and obeys an evolving $C^2$ hierarchy of rules. A global rule set determines combat attrition and reinforcement. Nonlinear feedback exists among combatants (measure $\rightarrow$ countermeasure $\rightarrow$ countercountermeasure $\rightarrow$ ...) and between combatants and the environment.

Note that this approach is similar in spirit to a cellular automaton (CA) model (see page 81) but augments the conventional CA framework in three ways: (1) evolution proceeds not according to a fixed set of rules, but to a set of rules that adaptively evolves over time; (2) individual states of cells (or combatants) do not just respond to local information, but are capable of non-local information (via an embedded $C^2$ topology) and command hierarchy; and (3) global rule (i.e. command) strategies are evolved via a genetic algorithm (orders pumped down echelon are based on evolved strategies played out on possible imprecise mental maps of local and/or global commanders).

Insofar as complex adaptive systems can be regarded as being essentially open-ended problem-solvers, their lifeblood consists mostly
of novelty. The ability of a complex adaptive system to survive and evolve in a constantly changing environment is determined by its ability to continually find - either by chance, or experience, or more typically both - insightful new strategies to increase its overall "fitness" (which is, of course, a constantly changing function in time).

Figure 38. Field-of-view of a single ISAACA

Military campaigns likewise depend on the creative leadership of their commanders, success or failure often hinging either on the brilliant tactic conceived in the heat of combat or the mediocre one issued in its place.

To be realistic, such novelty must not consist solely of a randomly selected option from a main-options list - which is a common approach taken by conventional warfare models - but must at least have the possibility of being as genuinely unanticipated in the model as it often is on a real battlefield. To this end, each command-agent (and to a somewhat more limited extent, each ISAACA) must possess both a memory and an internal anticipatory mechanism which it uses to select the optimal tactic and/or strategy from among a set of predicted outcomes. This is an important point: except for doctrine and the historical lessons of warfare, the super-set of tactics must not be hard-wired in.

Such local rule-based agent-simulations are well suited for

- studying the general efficacy of combat doctrine and tactics
- exploring emergent properties and/or other "novel" behaviors arising from low-level rules (even doctrine if it is well encoded\textsuperscript{20})

\textsuperscript{20} It is an intriguing speculation that doctrine as a whole may contain
• capturing universal patterns of combat behavior by focusing on a reduced set of critical drivers

• suggesting likelihood of possible outcomes as a function of initial conditions

• use as training tools along the lines of some commercially available agent-based "games" such as SimCity, SimFarm and SimLife [325]

• providing near-real-time tactical decision aids by providing a "natural selection" (via genetic algorithms; see page 93) of superior tactics and/or strategies for a given combat situation

• giving an intuitive "feel" for how and/or why unanticipated events occur on the battlefield, and to what extent the overall process is shaped by such events

Figure 39. Schematic representation of a ISAACA simulation

Ideally, one would hope to find universal patterns of behavior and/or tactics and strategies that are independent of the details of the makeup of individual ISAACAs.

both desirable and undesirable latent patterns that emerge only when allowed to "flow" through a system of elementary agents. An agent-based model of combat may provide an ideal simulation environment in which to explore such possibilities.
Agent-based simulations ought not be used either to predict real battlefield outcomes or to provide a realistic simulation of combat. While commercial networkable 3D virtual-reality games such as DOOM\(^21\) are much better suited to providing a virtual combat environment for training purposes, agent-based simulations are designed to help understand the basic processes that take place on the battlefield. It is not realism, for its own sake, that agent-based simulations are after, but rather a realistic understanding of the drivers (read: interactivity, decision-making capability, adaptability, and so on) behind what is really happening.

**Tactics and strategy evolution using genetic algorithms**

Genetic algorithms have been shown to be powerful tools for general combinatorial optimization search problems; see page 98. One obvious application of genetic algorithms that has immediately found its way into the artificial life research community involves their use as sources of the "adaptive intelligence" of adaptive autonomous agents in an agent-based simulation. A related application that is of particular interest to the military strategist and/or battlefield commander, is that of direct strategy and/or tactics development.

Figure 40. Schematic representation of a strategy landscape

\[^{21}\text{Id Software, World-Wide Web URL link = http://www.idsoftware.com.}\]
Virr, Fairley and Yates [310], for example, have proposed using a
genetic algorithm as the foundation of an Automated Decision
Support System (ADSS). Carter, et. al, [42] suggest using
 genetic-algorithm-derived strategies for "smart tanks." Crowston [61]
uses genetic algorithms to search for alternative organizational forms,
which has potential applications for rethinking the optimality of
military command and control structures.

Figure 40 shows a schematic representation of what might be called a
"strategy landscape." The strategy landscape represents the space of all
possible global strategies that can be followed in a given scenario.
Generally speaking, a genetic-algorithm-based tactics- or strategy-
"optimizer" consists of an evolutionary search of this landscape for
high-pay-off strategies using whatever local information is available to a
combatant. The shape of the landscape is determined by the fitness
measure that is assigned to various tactics and/or strategies. It also
changes dynamically in time, as it reorients to the actual search path
that is being traversed.

Time-series analysis

Time-series analysis deals with the reconstruction of any underlying
attractors, or regularities, of a system from experimental data
describing a system's behavior; see page 57. Techniques developed
from the study of nonlinear dynamical systems and complex systems
theory provide powerful tools whereby information about any
underlying regularities and patterns in data can often be uncovered.
Moreover, these techniques do not require knowledge of the actual
underlying dynamics; the dynamics can be approximated directly from
the data. These techniques provide, among other things, the ability to
make short- (and sometimes long-) term predictions of trends in a
system's behavior, even in systems that are chaotic.

Relativistic information

Relativistic information theory is a concept introduced by Jumarie
and has been suggested as a possible formalism for describing certain
aspects of military command and control processes by Woodcock and
Dockery [77]. The basic idea is that a generalized entropy is endowed
with four components, so that it is equivalent to a four-vector and may
be transformed by a Lorentz transformation (as in relativity theory).
These four components consists of: (1) the external entropy of the
environment (H_e), (2) the internal entropy of the system (H_s), (3) system
goals, and (4) the internal transformation potential, which measures the

\[ H = H_e + H_s + H_g + H_T \]

\[ Jumarie, G., "A relativistic information theory model for general
systems Lorentz transformation of organizability and structural entropy,"\nInternational Journal of Systems Science, Volume 6, 1975, 865-886.\]
efficiency of the system’s internal information transformation. An additional factor, called the *organizability*, plays the role of "velocity." Woodcock and Dockery show that it is possible to use relativistic information theory to compare the relative command and control system response of two command structures to the world around them. The quantity of interest is $\frac{dH_i}{dH_0}$, or the rate of change of the internal information environment with respect to changes in the surrounding environment.

**Exploitation of Characteristic Time Scales**

A fundamental property of nonlinear systems is that they generally react most sensitively to a special class of aperiodic forces. Typically, the characteristic time scales of the optimal driving force match at all times the characteristic time scales of the system. In some cases the optimal driving force as well as the resulting dynamics are similar to the transients of the unperturbed system.  

The information processing in complex adaptive systems and the general sensitivity of all nonlinear dynamical systems to certain classes of aperiodic driving forces are both potentially exploitable features. Recall that one of the distinguishing characteristics of complex systems is their information processing capability. Agents in complex adaptive systems continually sense and collect information about their environment. They then base their response to this information by using internal models of the system, possibly encoding and storing data about novel situations for use at a later time. According to the *edge-of-chaos* idea (see page 76), the closer a system is to the edge-of-chaos – neither too ordered nor too chaotic – the better it is able to adapt to changing conditions. In Kauffman’s words, "organisms sense, classify, and act upon their worlds. In a phrase, organisms have internal models of their worlds which compress information and allow action...Such action requires that the world be sufficiently stable that the organism is able to adapt to it. Were worlds chaotic on the time scale of practical action, organisms would be hard pressed to cope."  

Now compare this state-of-affairs with Retired USAF Colonel John Boyd’s *Observe-Orient-Decide-Act* (OODA) loop. In Boyd’s model, a system responds to an event (or information) by first observing it, then considering possible ways in which to act on it, deciding on a particular course of action and then acting. From a military standpoint, both friendly and enemy forces continuously cycle through this OODA process. The objective on either side is to do this more rapidly than the enemy; the idea being that if you can beat the enemy to the "punch" you can disrupt the enemy’s ability to maintain coherence in a

---

23 A. Hubler, "Modeling and control of complex systems: paradigms and applications," pages 5-65 in [82].

24 Page 232 in reference [171].
changing environment. One can also imagine exploiting the relative phase relationship between friendly and enemy positions within the OODA loop. For example, by carefully timing certain actions, one can effectively slow an enemy's battle-tempo by locking the enemy into a perpetual Orient-Orient mode.
Summary and Conclusion

This paper provides the basic theoretical framework and mathematical background necessary to intelligently discuss the ideas of nonlinear dynamics and complex systems theory and how they might apply to land warfare issues. Part II of this study, to be delivered to sponsor for review 1 July, 1996, will consist of a detailed assessment of the general applicability of complexity theory to the representation of land warfare.

Overall, the paper provides four separate levels of discussion:

- **Basic Concepts.** The first level consists of a discussion of the basic concepts of nonlinear dynamics and complex systems theory, including nonlinearity, chaos, phase space, attractors, fractals, predictability, etc., and thus provides a working technical vocabulary.

- **Mathematical Tools.** The second level consists of a discussion of specific mathematical tools that can be applied to the study of complex systems in general, such as Poincare plots, Lyapunov exponents, genetic algorithms, etc.

- **Basic Lessons Learned.** The third level consists of a discussion of basic lessons learned from both nonlinear dynamics and complex systems theory.

- **Possible Applications.** The fourth level consists of an introductory survey of possible applications of the tools and concepts of complexity theory to land warfare. This last level is preliminary and is intended only to "plant a few seeds" for an in-depth analysis in part II of this study.

Basic Concepts

A quick-reference glossary of all of the terms and basic concepts appearing in the main text are given in Appendix B.

Mathematical Tools

Below is a partial summary of the mathematical tools discussed in earlier sections:
• **Qualitative Characterization of Chaos.** Four qualitative methods for verifying the presence of chaos in a system were discussed. These included looking at the system's *time-dependent behavior*, using a *Poincare plot* to reduce the dimensionsality, calculating the *autocorrelation function* and observing the *power spectrum* for the system.

• **Quantitative Characterization of Chaos.** Three sets of quantitative measures of chaos were introduced, including *Lyapunov exponents* (that measure the exponential divergence of initially nearby trajectories), *generalized fractal dimensions* (that, roughly speaking, measure the minimum number of variables needed to specify a chaotic attractor), and the *Kolmogorov-Sinai entropy* (that measures the rate of information gain per unit time in observing a chaotic system).

• **Time-Delayed Embedding.** The embedding technique is a method for reconstructing a state space from time-series data. It assumes that if the embedding dimension is large enough, the behavior of whatever system is responsible for generating the data can be described by a finite dimensional attractor. Its main strength lies in providing detailed information about the behavior of degrees-of-freedom other than the ones that are directly observed.

• **Chaotic Control.** Chaotic control exploits the fact that chaotic systems exhibit sensitivity to initial conditions to stabilize regular dynamical behaviors and thereby effectively "direct" chaotic trajectories to a desired state.

• **Cellular Automata.** Cellular automata are a class of spatially and temporally discrete, deterministic dynamical systems characterized by local interaction and an inherently parallel evolution. They serve as prototypical mathematical models of complex systems, and appear to capture many essential features of complex self-organizing cooperative behavior observed in real systems.

• **Genetic Algorithms.** Genetic algorithms are a class of heuristic search methods and computational models of adaptation and evolution based on natural selection. Genetic algorithms mimic and exploit the genetic dynamics underlying natural evolution to search for optimal solutions of general combinatorial optimization problems. This very powerful tool is used frequently as the backbone of many artificial life studies.

• **Agent-Based Simulations.** Agent-based simulations of complex adaptive systems are predicated on the idea that the global
behavior of a complex system derives entirely from the low-level interactions among its constituent agents. By relating an individual constituent of a complex adaptive system to an agent, one can simulate a real system by an artificial world populated by interacting processes. Agent-based simulations are particularly adept at representing real-world systems composed of individuals that have a large space of complex decisions and/or behaviors to choose from.

- **Swarm.** Swarm (currently under development at the Santa Fe Institute) is a multi-agent simulation platform for the study of complex adaptive systems. The goal of the Swarm project is to provide the complex systems theory research community with a fully general-purpose artificial-life simulator. Swarm has been intentionally designed to include as few ad-hoc assumptions about the design of a complex system as possible, so as to provide a convenient, reliable and standardized set of software tools that can be tailored by researchers to specific systems.

- **Neural Networks.** Neural nets represent a radical new approach to computational problem solving. Their *bottom-up* methodology stands in stark contrast to traditional *top-down* approach to artificial intelligence (AI). The approach is inspired by such basic skills of the human brain as its ability to continue functioning with noisy and/or incomplete information, its robustness or fault tolerance, its adaptability to changing environments by learning, etc. Neural nets attempt to mimic and exploit the parallel processing capability of the human brain in order to deal with precisely the kinds of problems that the human brain itself is well adapted for; in particular, pattern recognition.

**Basic Lessons Learned**

**Nonlinear Dynamics**

The fundamental lesson of nonlinear dynamics is that a dynamical system does not have to be "complex" or to be described by a large set of equations, in order for the system to exhibit chaos – all that is needed is three or more variables and some embedded nonlinearity.

Among the basic lessons of nonlinear dynamics and chaos that are of particular relevance to the decision maker are...
• **Short Term Predictions.** Given sufficient data, time series analysis permits one to make short-term predictions about a system's behavior, even if the system is chaotic. Moreover, these prediction can be made even when the underlying dynamics is not known.

• **Long-term Trends.** If the attractors of a system are known or can be approximated (say, from available historical time series data), long-term trends can be predicted. Knowledge about visitation frequencies of points on an attractor provides insight into the probabilities of various possible outcomes. Lyapunov exponents quantify the limits of predictability.

• **Qualitative Understanding of the Battlefield.** The information dimension can be used to estimate the minimum number of variables needed to describe a system. Moreover, if a system can be shown to have a small non-integer dimension, it is probable that the underlying dynamics are due to nonlinearities and are not random.

**Complex Systems**

The fundamental lesson of complex systems theory is that complex behavior is usually an emergent self-organized phenomenon built upon the aggregate behavior of very many nonlinearly interacting "simple" components.

The most important points to remember are...

• **Nonlinearity.** Without nonlinear interactions there can be no deterministic chaos in simple systems and no complex behavior in complex systems. Moreover, nonlinear systems appear to be much more pervasive than linear systems. By virtue of nonlinearity, the behavior of the "whole" is not just a simple aggregate of the constituent "parts."

• **Interconnectivity.** How the parts of a complex system are interconnected is just as important as what those parts are and what does parts do.

• **Context/Wholeness.** The effect that parts have on the remainder of the system – literally, how those parts are defined within the complex system – is determined by the context of the whole within which those parts exist.

• **Process.** A complex system embodies process, a ceaseless search for a better "solution" for an ill-defined, amorphous ever
receding "problem." In Zen-like fashion, you can say that the harder one tries to pin-down the behavior of a complex system with some static measure, the further one is from understanding what the complex system is really doing.

- **Adaptability.** The essence of a complex adaptive system is that its constituent parts are not Newtonian "billiards" that react blindly (but in well-defined fashion) to the world around them, but are instead endowed with an ability to *sense, learn from,* and *adapt* to their environment as they and the environment both evolve in time.

- **Emergence.** Perhaps the central concept of complex systems theory is that high-level behaviors emerge naturally out a brewing soup of low-level interactions. A flock of birds does not need a central direction to behave in an apparently orchestrated manner. The lesson is that where there is an assemblage of very many nonlinearly interacting parts, there is a good possibility of emergent behaviors on higher levels than those defining the underlying interactions. Moreover, such emergent behavior can appear on multiple spatial and temporal levels.

**Possible Applications of Complexity Theory to Land Warfare**

The last section of this paper outlined a few preliminary musings on the applicability of nonlinear dynamics and complex systems theory to the understanding and/or representation of land warfare. An in-depth analysis and discussion of the ideas presented in that section will be provided in Part II of this report.

The fundamental question is *"What does complexity theory tell us about land warfare?"* The last section provided a framework for a possible answer to this question by focusing on (1) four levels of applicability of complexity theory (ranging from general tools, to specific simulation laboratories, to high-level properties of specific systems to universal behaviors), (2) four levels of land warfare (tactical, operational, strategic and general strategic), and (3) three levels on which complexity theory can be applied to land warfare:

- **Metaphor Level.** This level consists of constructing and elaborating upon similar sounding words and images that most strongly suggest a "philosophical resonance" between behaviors of complex systems and certain aspects of what happens on a battlefield.
Pragmatic and/or Experimental Level. This is the level on which real-world data is mined to confirm or deny that there is more to a possible connection between complexity theory and land warfare than mere "philosophical resonance" alone.

Direct Application Level. This is the "workhorse" level on which specific methodology borrowed from complexity theory is applied directly to specific issues and problems of land warfare.

There is also the possibility that complexity theory might shed some light on how battlefields may be configured (or compelled to self-organize) to achieve a maximum adaptability to a changing environment.

The remainder of the last section provided a brief overview of some specific applications:

Chaos in combat and arms-race models. A fundamental lesson of nonlinear dynamics theory is that one can almost always expect to find some manifestation of chaos whenever nonlinearities are present in the underlying dynamics of a model. This fundamental lesson has potentially significant implications for even the simplest combat models. Several instances of chaos in simple combat models were cited, including work by Miller and Sulcoski [217], Dockery and Woodcock [77] and a recent RAND study [73].

Generalizations of Lanchester's Equations. Generalizations include an analytical extension to allow feedback between movement and attrition (Epstein, [90]), an exploitation of the analogy between Lanchester's equations and the Lottka-Voltera equations describing natural ecologies, and partial, fuzzy and stochastic differential equations.

Demonstration of chaos in war using historical data. Tagarev, et. al.'s work was cited as providing evidence of chaos in tactical, operational and strategic levels of military activity.

Combat simulation using cellular automata and adaptive autonomous agents. Woodcock, Cobb and Dockery's cellular automata model of combat is cited [323], as well as a more elaborate adaptive autonomous agent model in which individual combatants are equipped with (1) a default rule set, (2) goals directing behavior, (3) sensors to generate a map of the environment, and (4) an internal mechanism to adaptively and selectively alter behavior over time. Reasons for studying such models are discussed.
• **Tactics and strategy evolution using genetic algorithms.** It is suggested that genetic algorithms be used to develop strategy and/or tactics. Genetic-algorithm-based tactics or strategy-"optimizers" would consist of an evolutionary search of a "strategy landscape" for high-pay-off strategies using whatever local information is made available to a combatant.

• **Intelligent software agents.** Software agents are essentially sophisticated personal assistants. They act as intermediaries between the interests of the user and the global information pool with which the user has traditionally dealt directly. From a military standpoint, intelligent software agents can be used for *adaptive information filtering and integration* and as *tactical picture agents*, scouring and putting order on the amorphous flood of battlefield and intelligence data.

• **Time-series analysis.** Techniques developed from nonlinear dynamics and complex systems theory provide powerful tools with which underlying regularities and patterns in data can often be uncovered.

• **Exploitation of characteristic time scales of a combat.** A fundamental property of nonlinear systems is that they generally react most sensitively to a special class of aperiodic forces. An analogy between John Boyd’s *Observe-Orient-Choose-Act* (OODA) loop and information processing at the *edge-of-chaos* in complex systems suggests ways of interfering with an enemy’s OODA "timing" and thereby disrupting the enemy’s ability to maintain a coherence in a changing environment.
Appendix A: World Wide Web Nonlinear Dynamics and Complex Systems Theory Resources

Appendix A provides both a brief subject-sorted listing of information sources currently available on the World Wide Web (WWW) and an unsorted but much more extensive alphabetized listing in HTML-format.

Subject-Sorted WWW Link Listing

In this section, a total of 91 WWW Universal Resource Locator (URL) links are sorted into the following 16 categories:

- General Sources
- Artificial Intelligence
- Artificial Life
- Artificial Life Simulation and Research Groups
- Autonomous Agents
- Cellular Automata
- Chaos
- Fractals
- Fuzzy Logic
- Genetic Algorithms
- Genetic Programming
- Intelligent Software Agents
- Neural Nets
- Nonlinear Dynamics
- Software
- Time-Series Analysis
Appendix A

General Sources

- Santa Fe Institute: http://www.santafe.edu/
- Complex Systems Research at the Beckman Institute: http://www.ccsr.uiuc.edu/
- The Chaos Network, applications of chaos theory to the social sciences: http://www-cse.ucsd.edu:80/users/rik/
- Bibliography of Measures of Complexity: http://www.fmb.mmu.ac.uk/~bruce/combib
- Principia Cybernetica: http://pespmc1.vub.ac.be/

Artificial Intelligence

- International's Artificial Intelligence Center: http://www.ai.sri.com:80/aic/
Appendix A

- Journal of Artificial Intelligence Research:
  http://www.cs.washington.edu/research/jair/home.html

- Artificial Intelligence (Georgia Institute of Technology):
  http://www.cc.gatech.edu/cogsci/ai.html

- Distributed Artificial Intelligence Laboratory (UMass):
  http://dis.cs.umass.edu/

Artificial Life

- Adaptive Systems and Artificial Life:
  http://doradus.einet.net/galaxy/Engineering-and-Technology.html

- Philosophy of Artificial Life Bibliography:
  http://mugwump.ucsd.edu/bkeeley/work-stuff/Alife_Bib.html

- Artificial Life Bibliography:
  ftp://cognet.ucla.edu/pub/alife/papers/alife.bib.gz

- A Semi-annotated Artificial Life Bibliography:
  http://www.cogs.susx.ac.uk/users/ezequiel/alife-page/alife.html

- Artificial Life Digest:
  http://www.cogs.susx.ac.uk/users/ezequiel/alife-page_complexity.html

- Fundamental Algorithms of Artificial Life:
  http://alife.santafe.edu/alife/topics/simulators/dret/dret.html

Artificial Life Simulation and Research Groups

- Links to various Artificial Life Groups:
  http://www.krl.caltech.edu/brown/AL-groups.html

- Autonomous Agents/Alife Group at MIT:
  http://agents.www.media.mit.edu/groups/agents/

- Latent Energy Environments Project:
  http://www-cse.ucsd.edu:80/users/fil/
Appendix A

- The Avida Artificial Life group:
  http://www.krl.caltech.edu/avida/Avida.html

- Distributed Intelligent Agents Group:
  http://ai.eecs.umich.edu/diag/homepage.html

- Autonomous Agents Research Group:
  file://alpha.ces.cwru.edu/pub/agents/home/html

**Autonomous Agents**

- Autonomous Agents Group at MIT:
  http://agents.www.media.mit.edu/groups/agents/

- Autonomous Agents Group at Case Western Reserve University:
  http://yugggoth.ces.cwru.edu/

- Research on Autonomous Agents at Stanford:
  4-nilsson.html

- Software Agents Mailing List:
  http://www.sml.com/research/tcl/lists/AGENTS/index.html#163

- Autonomous Agents Research at Buffalo:
  http://www.cs.buffalo.edu/~jweber/autoagent.html

- Distributed Intelligent Agents Group at the University of Michigan:
  http://ai.eecs.umich.edu/diag/homepage.html

**Cellular Automata**

- Cellular Automata Frequently Asked Questions (FAQ):
  http://alife.santafe.edu/alife/topics/cas/ca-faq/ca-faq.html

- The Cellular Automata Simulation System:
  http://www.cs.runet.edu/~dana/ca/cellular.html

- Cellular Automata Web:
  http://alife.santafe.edu/alife/topics/ca/caWeb
Appendix A

- Cellular Automata Bibliography Database:
  http://www ima.umn.edu/bibtex/ca.bib

- Usenet Cellular Automata Newsgroup:
  news:comp.theory.cell-automata

- Evolving Cellular Automata:
  http://www.santafe.edu/projects/evca/

Chaos

- Applied Chaos Laboratory at Georgia Tech:
  http://acll.physics.gatech.edu/aclhome.html

- Chaos Bibliography:
  http://www.uni-mainz.de/FB/Physik/Chaos/services.html

- Chaos e-Print Archive at Los Alamos:
  http://xxx.lanl.gov/archive/choa-dyn/

- Chaos Group at the University of Maryland at College Park:
  http://www-chaos.umd.edu/

- The Chaosgruppe (Munchen):
  http://www.nonlin.tu-muenchen.de/chaos/chaos_e.htm1

Fractals

- Exploring Chaos and Fractals (MIT):

- Fractals FAQ:
  http://www.cis.ohio-state.edu/hypertext/faq/usenet/fractal-faq/faq.html

- Fractal Images:
  http://www.acm.uiuc.edu:80/rml/Gifs/Fractal/

- Fractal Pictures and Animations:
  http://www.cnam.fr/fractals.html
Appendix A

- Fractal Explorer:  
  http://www.vis.colostate.edu/~user1209/fractals/index.html

- Fractal Database: http://spanky.triumf.ca/

**Fuzzy Logic**

- Fuzzy Logic Archive:  

- Fuzzy Logic FAQ:  

- Fuzzy Logic Repository:  
  ftp://ntia.its.bldrdoc.gov/pub/fuzzy

**Genetic Algorithms**

- Illinois Genetic Algorithm Repository: 
  http://gal4.ge.uiuc.edu/

- Genetic Algorithms FAQ:  

- Interactive Genetic Art:  
  http://robocop.modmath.cs.cmu.edu:8001/htbin/mjwgenformI

- Genetic Music: http://nmt.edu/~jefu/notes/notes.html

- Genetic Algorithm Digest Archives:  

- Genetic Algorithms Tutorial:  
  ftp://129.82.102.183/pub/TechReports/1993/tr-103.ps.Z
Appendix A

**Genetic Programming**

- Genetic Programming FAQ:  
  http://wwwhost.cc.utexas.edu/cc/staff/mccoy/gp/FAQ-toc.html

- Genetic Programming Tutorial:  
  http://dcpul.cs.york.ac.uk:6666/mark/top_ga.html

- Genetic Programming in C++ FAQ:  
  http://www.salford.ac.uk/docs/depts/eee/gpfaq.html

- Genetic Programming Source at UCL:  
  http://www.cs.ucl.ac.uk/intelligent_systems/genetic_programming.html

- Genetic Programming Bibliography:  
  ftp://cs.ucl.ac.uk/genetic/biblio/

- Genetic Programming FTP site:  

**Intelligent Software Agents**

- Software Agents:  
  http://hitchhiker.space.lockheed.com/pub/AGENTS/htdocs/agent-home.html

- Intelligent Software Agents Resources:  
  http://retriever.cs.umbc.edu:80/agents/

- Intelligent Agents Mailing List (by thread):  
  http://www.smli.com/research/tcl/lists/AGENTS/index.html#163

- MIT Media Lab: http://www.media.mit.edu/

- Intelligent Software Agents (University of Maryland Baltimore County):  
  http://www.cs.umbc.edu/agents/

**Neural Nets**

- A Basic Introduction To Neural Networks:  
  http://ice.gis.uiuc.edu/Neural/neural.html
Appendix A

- An Introduction to Neural Nets:
  http://www.mindspring.com/~zsol/nintro.html

- Neural Networks FAQ:
  http://www.eeb.ele.tue.nl/neural/neural_FAQ.html

- Collection of Neural Net Bibliographies:
  http://glimpse.cs.arizona.edu:1994/bib/Neural/

- IEEE Neural Networks Council:

- International Neural Network Society:
  http://sharp.bu.edu/inns/

**Nonlinear Dynamics**

- Nonlinear Dynamics Archive: ftp://lyapunov.ucsd.edu/pub

- Nonlinear Dynamics e-print Archive at Los Alamos:
  http://xyz.lanl.gov/

- Nonlinear Dynamics and Topological Time Series:
  http://t13.lanl.gov/~nxt/intro.html

- Institute of Nonlinear Science at UC San Diego:
  http://inls.ucsd.edu/

- Nonlinear Dynamics at UC Santa Cruz:
  http://noether.ucsc.edu/groups/nonlinear/research.html

- Nonlinear Dynamics Sites:
  http://www.ucl.ac.uk/~ucesjph/resources/uk.html

**Software**

- Artificial Life Software at Santa Fe Institute:
  http://alife.santafe.edu/alife/software/

- Artificial Life Software Repository:
  http://www.cs.cmu.edu/afs/cs/project/ai-repository/ai/areas/alife/systems/0.html
Appendix A

- Cellular Automata Simulator for PC Windows:
  fe/rudy-rucker/

- Complex Systems Software Repository at Australian National

- WinLife (a PC Windows implementation John Conway's Life rule):
  fe/packages/winlife

- WinCA (a fast PC Windows simulator):
  fe/packages/winca

- PC Windows implementation of Craig Reynolds "Boids":
  fe/packages/boids/

**Time Series Analysis**

- Nonlinear Dynamics and Topological Time Series Analysis Archive:
  http://t13.lanl.gov/~nxt/intro.html
Appendix A

Alphabetized WWW Link Listing in HTML format

<p>
<center>ch1:Complex Systems Links</center>
</p>

<p>
<A NAME="aa">
</A CENTER>
</p>

<p>
<HR SIZE=7MM>
</p>

<p>
<CENTER><H2>
<A HREF="#A">A</A>
<A HREF="#B">B</A>
<A HREF="#C">C</A>
<A HREF="#D">D</A>
<A HREF="#E">E</A>
<A HREF="#F">F</A>
<A HREF="#G">G</A>
<A HREF="#H">H</A>
<A HREF="#I">I</A>
<A HREF="#J">J</A>
<A HREF="#K">K</A>
<A HREF="#L">L</A>
<A HREF="#M">M</A>
<A HREF="#N">N</A>
<A HREF="#O">O</A>
<A HREF="#P">P</A>
<A HREF="#Q">Q</A>
<A HREF="#R">R</A>
<A HREF="#S">S</A>
<A HREF="#T">T</A>
<A HREF="#U">U</A>
<A HREF="#V">V</A>
<A HREF="#W">W</A>
<A HREF="#X">X</A>
<A HREF="#Y">Y</A>
<A HREF="#Z">Z</A>
</H2></CENTER>
</p>

<p>
<HR SIZE=7MM>
</p>

<p>
<DT><DD><H1><A NAME="A">A</A></H1>
</DT><DD><OL>
<DT><DD><LI><A HREF="http://ice.gis.uiuc.edu/Neural/neural.html">A Basic Introduction To Neural Networks</A>
</DT><DD><LI><A HREF="http://www.cogs.susx.ac.uk/users/ezquiel/alife-page/alife.html">A Semi-annotated Artificial Life Bibliography</A>
</DT><DD><LI><A HREF="http://www.krl.caltech.edu/~adami/">Chris Adami's Homepage</A>
</DT><DD><LI><A HREF="http://www-ksl.stanford.edu/people/bhr/index.html">Adaptive Intelligent Systems</A>
(Stanford University)
</DT><DD><LI><A HREF="http://iserv.iki.fkibi.hu/adaptlab.html">Adaptive Systems Laboratory</A>
</DT><DD><LI><A HREF="http://doradus.cinet.net/galaxy/Engineering-and-Technology.html">Adaptive Systems and Artificial Life</A>
</DT><DD><LI><A HREF="http://borneo.gmd.de/AS/pages/as.html">Adaptive Systems Research Group</A>
German National Research Center for Computer Science
</DT><DD>
</OL>
</DD>
</DT>
</DD>

158
Appendix A

- Artificial Life Group (Iowa State University)
- Australian National University Bioinformatics
- Avida
- Artificial Life approaches with Mobile Fischertechnik Robots
- Aquarium
- Artificial-life bibliography (huge)
- A Semi-annotated Artificial Life Bibliography
- Artificial Life Games Homepage
- Artificial Life Groups
- Artificial Life Journal
- Artificial Life Lab
- Artificial Life Related Newsgroups Archive
- Artificial Life Page
- Artificial Life Online
- Artificial Life Homework
- Artificial Life On WWW (Italy)
- Artificial Life resources (by Patrick Tufts)
- Artificial Life resources (by Titus Brown)
- Artificial Life resources (Yahoo)
- Artificial Life Software (Santa Fe Institute)
- Artificial Life Software at CMU
- Online ALife Examples (Yahoo)
- Artificial Life and Complex Systems Catalogue (Y. Kanada)
- Artificial Life at HPLB
- Philosophy of Artificial Life Bibliography
- Artificial-Life Simulators and Their Applications
- Dave's A-Life Pages
- Artificial Life software packages
- Artificial Living Room
- a place to discuss evolutionary theory
- Artificial neural networks: a developing science
- Artificial Neural Network Research at NEC Research Institute, Princeton, NJ
- Artificial Painter: a combination of Genetic Algorithms and Neural Networks
- Artificial pets with real brains (Nick Turner's collection)
- Artificial Society Group in Japan
Appendix A

Association for Uncertainty in Artificial Intelligence

Auburn NonLinear Dynamics

Austen Center for Nonlinear Dynamics

Australian National University Bioinformatics

Austrian Research Institute for Artificial Intelligence: Neural Net Group

Automatrix CA Hardware (CAM-PC)

Autonomous Agents Group

Autonomous Agents Research Group (MIT)

Autonomous Agents Research Group (Case Western Reserve University)

Autonomous Agents Group (homepage at CWRU)

Autonomous Systems Group (Robots and Neurocomputing)

Avida

Avida Artificial Life group

Back to top of page......

Baldwin Effect

Baldwin Effect Bibliography

Bayesian Model-Based Learning Group

Behavioral Evolution Simulations and Tutorials

Bibliographies on Neural Networks

Biodiversity and Biological Collections

Biological Computation Project at University of Alberta, Canada

Biological time-series web page from Argentina

Biomorphic

Biomorphic (2)

The Alife Library

The bibliographic collection at Alife Online

A bibliography of readings on complex adaptive systems

Bibliography of Measures of Complexity

Biota.org

Boids

Boids for Windows

Bomb

Boston University Mathematics Department and Dynamical Systems Group
Appendix A

A HREF="http://robotics.stanford.edu/groups/bots/home.html">Bots Research Group</a> (Stanford University)

A HREF="http://synapse.cs.byu.edu/home.html">Brigham Young University (BYU) Neural Networks and Machine Learning Lab</a>

A HREF="http://www.ucl.ac.uk/~ucesjph/resources/uk.html">British Nonlinear Sites</a>

A href="http://www.ai.mit.edu/people/brooks/brooks.html">Rodney A. Brooks' Homepage</a>

A href="http://www.cns.brown.edu/ibns/">Brown University: Institute for Brain and Neural Systems</a>

Buenos Aires (Univ. of BA) Chuchi Server: Non-linear time series analysis</a>


A HREF="ftp://life.anu.edu.au/pub/complex_systems/ alife/">Bugworld</a> (2)

A HREF="http://prairienet.org/business/ptech/chaos.html">Business</a> and Chaos

A href="http://synapse.cs.byu.edu/home.html">BYU Neural Networks and Machine Learning Laboratory</a>

Back to top of this page......</A>/I>

C language Genetic Programming System</a>

C++ Multi-agent simulator</a> for Unix System by Renaud Cazoulat (France)

Calife</a> a 1D CA simulator by Rudy Rucker

Calife</a> a 2D CA simulator by Rudy Rucker

California (UC Santa Cruz): Nonlinear Dynamics</a>

California: UC at San Diego, the Institute for Nonlinear Science</a>

California (UC at Davis) Center for Neuroscience</a>

Caltech: Physics of Computation, Computation and Neural Systems Program</a>

Caltech: The Koch Lab</a>

Callahan's Life page</a>

CAM8</a> (MTT)

Carnegie-Mellon University: The Center for the Neural Basis of Cognition</a>

CATS</a> Chaos and Turbulence Studies Center at NBI

Michael Creutz's CA simulators for Xwindows</a>

A Cellular Automaton Tool

CA Web</a>

Cellsim</a> is a cellular automaton simulator by David Diebeler and Chris Langton

Cellular Automata</a>

Cellular Automata</a> by Juha Haataja

Cellular Automata Mailing List Archive</a>
Appendix A

<DT><DD><LI><A href="http://www.seas.upenn.edu/~ale/cplxsys.html">Complex Adaptive Systems</A><br/>(Overview)
<DT><DD><LI><A HREF="ftp://life.anu.edu.au/pub/complex_systems">Complex Systems ftp directory from ANU</A>
<DT><DD><LI><A HREF="http://bambi.ccs.fau.edu/ccs.html">Complex Systems</A><br>(Boca Raton)
<DT><DD><LI><A HREF="http://www.ccsr.uiuc.edu/">Complex Systems</A><br>(at Beckman Institute)
<DT><DD><LI><A href="http://pscs.physics.lsa.umich.edu/pscs.html">Complex Systems</A><br>(University of Michigan)
<DT><DD><LI><A href="http://alife.santafe.edu/alife/topics/cas/ca-faq/ca-faq.html">Comp.theory.cell-automata FAQ</A>
<DT><DD><LI><A href="newscomp.theory.cell-automata">Comp.theory.cell-automata</A><br>(Usenet newsgroup)
<DT><DD><LI><A href="http://www.cns.caltech.edu/">Computation & Neural Systems Program</A><br>(CNS)
<DT><DD><LI><A HREF="http://bcn.boulder.co.us/environment/Global/EnvTopics.html">Computational Biology</A><br>(at SDSC)
<DT><DD><LI><A HREF="http://golgi.harvard.edu/biopages/">Computational Biology</A><br>(at Harvard)
<DT><DD><LI><A HREF="http://www.cse.ucsc.edu/research/compbio/">Computational Biology</A><br>(at UC Santa Cruz)
<DT><DD><LI><A href="http://beowulf.uwaterloo.ca/">Computational Epistemology Lab</A><br>(University of Waterloo)
<DT><DD><LI><A href="http://www.cirl.uoregon.edu/">Computational Intelligence Research Laboratory</A><br>(University of Oregon)
<DT><DD><LI><A HREF="http://coli.uni-sb.de/info/cl_in_sb.index.html">Computational Linguistics</A><br>(at Saarbrucken)
<DT><DD><LI><A href="http://www.santafe.edu/projects/CompMech/">Computational Mechanics Group at Santa Fe</A>
<DT><DD><LI><A HREF="http://monet.physik.unibas.ch:80/thomas/index.html">Condensed</A><br>(matter systems studied at University of Basel (Switzerland)
<DT><DD><LI><A href="http://www.cs.cmu.edu/Web/Groups/CNBC/other/connectionists.html">Connectionists</A>
<DT><DD><LI><A HREF="http://cdps.cs.unh.edu/">Cooperative Distributed Problem Solving</A><br>(University of New Hampshire)
<DT><DD><LI><A HREF="ftp://soda.berkeley.edu/pub/corewar">Corewar</A><br>(Berkeley,ftp)
<DT><DD><LI><A href="ftp://ftp.Germany.EU.net/pub/research/softcomp/Alife/ak-dewdney/">Core Wars</A><br>(ftp)
<DT><DD><LI><A HREF="http://www.tc.cornell.edu/Research/Articles/MPS/DMS/Durrett/durrett.models.html">Cornell Theory Center Movies</A>
<DT><DD><LI><A HREF="http://www.batnet.com/quit/fha/cr/">Critters</A>
<DT><DD><LI><A HREF="http://www.santafe.edu/~jpc/">James Crutchfield's Homepage</A>
<DT><DD><LI><A HREF="http://alife.santafe.edu/pub/CURRICULA/">CURRICULA</A><br> Syllabus suggestions for courses on Artificial Life
<DT><DD><LI><A HREF="http://pespmc1.vub.ac.be/journals.html">Cybernetics and Systems Journal</A>
<DT><DD><LI><A HREF="http://pespmc1.vub.ac.be/CYBSYSTH.html">Cybernetics and Systems Theory</A>
</DL>
</DT><DD></DD>
</DL>
</HR>
Appendix A

[H1]<A NAME="D">D</A></H1>
</DL>
<OL>
<DT><DD><DL><A HREF="http://alife.santafe.edu/alife/software/ ddlab.html">Discrete Dynamics Lab</A></DL>
<DT><DD><DL><A HREF="http://euler.mcs.uitlusa.edu/~sandip/ sandip.html">Distributed Artificial Intelligence</A></DL>(Tufts University)
<DT><DD><DL><A HREF="http://disc.cs.umass.edu/">Distributed Artificial Intelligence Laboratory</A>(UMass)
<DT><DD><DL><A HREF="http://ai.eecs.umich.edu/diag/homepage.html">Distributed Intelligent Agents Group</A>(University of Michigan)
<DT><DD><DL><A HREF="http://www.iesd.auc.dk/general/DS/index.html">Distributed Systems</A>
<DT><DD><DL><A HREF="http://www.wmin.ac.uk/~cdva/">Dynamical Symmetries</A>
<DT><DD><DL><A HREF="ftp://parcftp.xerox.com/pub/dynamics/ dynamics.html">Dynamics of Computation</A>(Xerox Palo Alto Research Center, ftp)
<DT><DD><DL><A HREF="ftp://parcftp.xerox.com/pub/dynamics/ multiagent.html">Dynamics of Multiagent systems</A>
</DL>
</DD>
</DL>
</DL>
</A HREF="#aa">Back to top of this page.....</A></I>
</DL>
<DL><HR>
</DL>
</DL>
</A NAME="E">E</A></H1>
</DL>
</OL>
<OL>
<DT><DD><DL><A HREF="http://alife.santafe.edu/alife/software/echo.html">Eco</A>(a) is an ecological simulation system by Terry Jones and John Holland
<DT><DD><DL><A HREF="http://www.cs.runet.edu/~dana/ca/ cellular.html">Eckart's Cellular Automata Simulator</A>
<DT><DD><DL><A HREF="ftp://alife.santafe.edu/pub/USER-AREA/EC/">Evolutionary Computation Repository (ECORE)</A> at Santa Fe Institute
<DT><DD><DL><A HREF="ftp://ftp.dcs.warwick.ac.uk/pub/mirrors/EC/">ECORE</A> at The University of Warwick, UK
<DT><DD><DL><A HREF="ftp://ftp.krl.caltech.edu/pub/EC/">ECORE</A> at The California Institute of Technology
<DT><DD><DL><A HREF="ftp://ftp.cs.wayne.edu/pub/EC/">ECORE</A> at Wayne State University, Detroit
<DT><DD><DL><A HREF="ftp://ftp.uct.ac.za/pub/mirrors/EC/">ECORE</A> at The University of Capetown, South Africa
<DT><DD><DL><A HREF="http://www.cns.ed.ac.uk/">Edinburgh: University of Edinburgh Centre for Neural Systems</A>
<DT><DD><DL><A HREF="http://kant.irmkant.rmn.cnr.it/u/ gral/luigi/ lupacnrgames.html">Educational & Therapeutic Alife Games</A>(online review (Italy-Denmark))
<DT><DD><DL><A HREF="ftp://alife.santafe.edu/pub/USER-AREA/EC/">ENCORE</A>(ftp)
<DT><DD><DL><A HREF="ftp://ftp.dcs.warwick.ac.uk/pub/mirrors/EC/Welcome.html">ENCORE Evolutionary Computation Archive</A>(2)
<DT><DD><DL><A HREF="ftp://nagun.aee.uic.edu/">Engineering</A>Nonlinear Systems Group at UIUC
<DT><DD><DL><A HREF="ftp://essex.ac.uk/pub/robots">Essex Robotics FTP Directory</A>
<DT><DD><DL><A HREF="http://www.cc.duth.gr/~mboudour/nonlin.html">European A nonlinear archive and pointer to nonlinear and complex sites</A>
<DT><DD><DL><A HREF="http://www.santafe.edu/projects/evca/">Evolving Cellular Automata</A>
<DT><DD><DL><A HREF="HTTP://www.sepa.tudelft.nl/~afd ba/ evolu.html">Evolution, Complexity and Philosophy</A>
<DT><DD><DL><A HREF="http://lancet.mit.edu/ga/OtherSites.html">Evolutionary Algorithm Sites</A>
</A>
Appendix A

<DT><DD><LI><A HREF="http://www.dai.ed.ac.uk/groups/evalg/">Evolutionary Algorithms Group</A> at The University of Edinburgh, UK
<DT><DD><LI><A HREF="http://www.cogs.susx.ac.uk/lab/adapt/index.html">Evolutionary and Adaptive Systems at COGS</A>
<DT><DD><LI><A HREF="http://www.cs.wisc.edu/~smucker/EC.html">Evolutionary Computation and Artificial Life</A>
<DT><DD><LI><A HREF="http://www.mitpress.mit.edu/jrnlscatalog/evolution.html">Evolutionary Computation Journal</A>
<DT><DD><LI><A HREF="http://zen.btc.uwe.ac.uk/evol/index.html">Evolutionary Computing Group at UWE, Bristol</A>
<DT><DD><LI><A HREF="http://pespmc1.vub.ac.be/EVOLSYS.html">Evolutionary Systems</A> an exploratory paper
<DT><DD><LI><A HREF="http://www.santafe.edu/projects/evca/index.html">Evolving Cellular Automata (EVCA) Group</A> Santa Fe Institute
<DT><DD><LI><A HREF="http://www.batnet.com/quist/fha/">Evolving Software (Evolutionary programming), Critters, CyberChromes</A>

</DL>
</DL>
Appendix A

[Links and references to various websites and resources related to fractals, artificial life, fuzzy logic, genetic algorithms, and artificial intelligence, including specific websites, research groups, and tools.]
Appendix A

App.</>
Appendix A

Back to top of this page....</A></DL>

ICE Neural Nets Hot List</A>

IEEE Neural Networks Council</A>

IlliGAL Home Page (Univ. Illinois-Urbana Genetic Algorithms Lab)</A>

Illinois (UI at Chicago): Discrete Dynamical Systems</A>

Illinois: UI at Urbana-Champaign, Center for Complex Systems Research</A>

Illinois: UI at Urbana-Champaign, The Beckman Institute for Advanced Science and Technology</A>

Images of Chaos</A>

Index to GA and artificial life resources</A>

Index to Complex Systems resources</A>

Information Mechanics (MIT)</A>

Institut Dalle Molle d'Intelligence Artificielle Perceptive</A>

Institute for Nonlinear Science at UCSD</A>

Institute for Language Technology and Artificial Intelligence</A>

Istituto Dalle Molle di Studi sull'Intelligenza Artificiale</A>

Intelligent agents</A>

Intelligent Hybrid Systems</A>

Institute for Solid State Physics and Chaos Group, Budapest</A>

Intelligent Software Agents</A>

Intelligent Systems Group</A>

Interactive Genetic Art</A>

Interactive Genetic Movies</A>

Interface Agents</A>

International Institute for Applied Systems Analysis

International Interactive Genetic Art</A>

International Interactive Genetic Art II Exhibit</A>

International Institute for Applied Systems Analysis</A>

International Neural Network Society</A>

International Philosophical Preprint Exchange</A>

International Society for Adaptive Behavior</A>

International Workshop on Decentralized Intelligent and multi-agent systems</A>

Introduction to Systems Theory and Complexity</A>
Appendix A


<OL>
<DL><DT><DD><LI><A HREF="#aa">Back to top of this page.....</A></DD></LI></DL>

<DL><DT><DD><LI><A HREF="http://alife.santafe.edu/alife/software/jvn.html">JVN</A> An implementation of the John von Neumann Universal Constructor</LI></DL>

<DL><DT><DD><DD><DD><A HREF="#aa">Back to top of this page.....</A></DD></LI></DL>

<OL>
<DL><DT><DD><LI><A HREF="http://www3s.unice.fr/~om/khepera.html">Khepera Simulator</A> public domain C/C++ package for writing a controller for a mobile robot</LI></DL>
<DL><DT><DD><LI><A HREF="http://www.is.cs.utwente.nl:8080/~kbs/kbsgeneralpage.html">KBS (Knowledge Based Systems) group</A></LI></DL>
<DL><DT><DD><LI><A HREF="http://www.cs.umbc.edu/kqml">Knowledge Query and Manipulation Language</A> (University of Maryland, Baltimore)</LI></DL>
<DL><DT><DD><LI><A HREF="http://logic.stanford.edu/knowledge.html">Knowledge Sharing</A> (Stanford University)</LI></DL>
<DL><DT><DD><LI><A HREF="ftp://cochlea.hut.fi/pub/">Teuvo Kohonen's Self-Organizing Map and Learning Vector Quantization software</A></LI></DL>
<DL><DT><DD><LI><A HREF="http://www.ncrl.postech.ac.kr">Postech Lab in Nonlinear Science</A></LI></DL>

<DL><DT><DD><DD><DD><A HREF="#aa">Back to top of this page.....</A></DD></LI></DL>

<OL>
<DL><DT><DD><LI><A HREF="http://www.santafe.edu/~cgl/">Chris Langton's Homepage</A></LI></DL>
Appendix A

<table>
<thead>
<tr>
<th>URL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://isl.msu.edu/GA/">http://isl.msu.edu/GA/</a></td>
<td>Michigan State University Genetic Algorithms Research and Applications Group</td>
</tr>
<tr>
<td><a href="http://www.ai.mit.edu/">http://www.ai.mit.edu/</a></td>
<td>MIT Artificial Intelligence Laboratory</td>
</tr>
<tr>
<td><a href="http://www.media.mit.edu/">http://www.media.mit.edu/</a></td>
<td>MIT Media-Lab</td>
</tr>
<tr>
<td><a href="http://www.santafe.edu/~mm/">http://www.santafe.edu/~mm/</a></td>
<td>Melanie Mitchell's Homepage</td>
</tr>
<tr>
<td><a href="http://www.dai.ed.ac.uk/groups/mrg/MRG.html">http://www.dai.ed.ac.uk/groups/mrg/MRG.html</a></td>
<td>Mobile Robots Group at the University of Edinburgh</td>
</tr>
<tr>
<td><a href="http://www.fusebox.com/ch/morphs/docs.html">http://www.fusebox.com/ch/morphs/docs.html</a></td>
<td>Morphs Evolution Game</td>
</tr>
<tr>
<td><a href="http://www.cogs.susx.ac.uk/users/davec/pe.html">http://www.cogs.susx.ac.uk/users/davec/pe.html</a></td>
<td>Movies of evolved pursuit/evasion strategies</td>
</tr>
<tr>
<td><a href="http://http2.brunel.ac.uk:8080/~hsrkngr/NNcourse/">http://http2.brunel.ac.uk:8080/~hsrkngr/NNcourse/</a> entry.html</td>
<td>MSc Intelligent Systems Neural Nets course material</td>
</tr>
<tr>
<td><a href="http://www.cs.utexas.edu/users/mfk/index.html">http://www.cs.utexas.edu/users/mfk/index.html</a></td>
<td>Multifunctional Knowledge Base Group (University of Texas at Austin)</td>
</tr>
<tr>
<td><a href="http://pianchet.rutgers.edu/">http://pianchet.rutgers.edu/</a></td>
<td>Nanotechnology Archive</td>
</tr>
<tr>
<td><a href="http://www.lucifer.com/~sean/Nano.html">http://www.lucifer.com/~sean/Nano.html</a></td>
<td>Nanotechnology Links and Pointers</td>
</tr>
<tr>
<td><a href="http://fas.sfu.ca/cs/research/groups/NLL/">http://fas.sfu.ca/cs/research/groups/NLL/</a> toc.html</td>
<td>Natural Language Laboratory (Simon Fraser University)</td>
</tr>
<tr>
<td><a href="http://cl-www.dfki.uni-sb.de/cl/registry/">http://cl-www.dfki.uni-sb.de/cl/registry/</a> draft.html</td>
<td>Natural Language Software Registry (German Research Institute for AI in Saarbruecken)</td>
</tr>
<tr>
<td><a href="http://hebb.cis.uoguelph.ca/home/ns.html">http://hebb.cis.uoguelph.ca/home/ns.html</a></td>
<td>Natural Selection Research Group (University of Guelph)</td>
</tr>
<tr>
<td><a href="http://www.cs.cmu.edu/afs/cs.cmu.edu/project/alv/member/www/navlab_home_page.html">http://www.cs.cmu.edu/afs/cs.cmu.edu/project/alv/member/www/navlab_home_page.html</a></td>
<td>NAVLAB (Carnegie Mellon University)</td>
</tr>
<tr>
<td><a href="http://web.mit.edu/~redingtn/www/nextadv/">http://web.mit.edu/~redingtn/www/nextadv/</a></td>
<td>The Net Advance of Physics is a journal/encyclopedia covering all areas of physics</td>
</tr>
<tr>
<td><a href="http://alife.santafe.edu/alife/software/">http://alife.santafe.edu/alife/software/</a> netlife.html</td>
<td>Netlife Evolving neural nets in an environment</td>
</tr>
<tr>
<td><a href="http://www.mech.gla.ac.uk/~nactftp/nact.html">http://www.mech.gla.ac.uk/~nactftp/nact.html</a></td>
<td>Neural Adaptive Control Technology</td>
</tr>
<tr>
<td><a href="http://www.dcs.shef.ac.uk/research/groups/ainn/">http://www.dcs.shef.ac.uk/research/groups/ainn/</a></td>
<td>Neural Computing</td>
</tr>
<tr>
<td><a href="http://synap.neuro.sfc.keio.ac.jp/">http://synap.neuro.sfc.keio.ac.jp/</a></td>
<td>Neural Computing Center (Keio University, Fujisawa, Japan)</td>
</tr>
<tr>
<td><a href="http://neural-server.aston.ac.uk/">http://neural-server.aston.ac.uk/</a></td>
<td>Neural Computing Research Group</td>
</tr>
</tbody>
</table>
Appendix A

- A Basic Introduction To Neural Networks
- Neural Networks at the Molecular Science Research Center
- Neural Networks at the Pacific Northwest Laboratory
- Neural Networks (CBMR): List
- Neural Networks at Los Alamos
- Neural Networks Using Genetic Algorithms
- Neural Network Software
- Neurocomputing WWW site British Department of Trade and Industry
- NEuroNet
- Neuroprose Archives at Ohio State University
- Neurosciences Internet Resource Guide
- Newsgroup: comp.theory.cell-automata
- NIA&R Artificial Intelligence & Robotics Group
- Niels Bohr Institute CATS - Center for Chaos and Turbulence Studies
- Nonlinear Dynamics (University of Mainz)
- Nonlinear dynamics and complexity
- Nonlinear dynamics and patterns
- Nonlinear Dynamics Site List
- Nonlinear Dynamics and Topological Time Series Analysis Archive
- Nonlinear Dynamics Archive
- Nonlinear science ePrint Archive
- Nonlinear Science Today (Springer-Verlag)
- Nonlinearity and Complexity Home Page

Back to top of this page.....

Back to top of this page.....

173
Appendix A

Research on Autonomous Agents</a> (Stanford)

Resource Guide to Complex Systems on the Net</a>

Craig W. Reynolds' Homepage</a>

Robotics Simulation Program</a> (ftp)

Robotics FAQ</a>

Robotics Internet Resources Page</a>

Robotic Systems & Advanced Computer Technology Section</a> (JPL)

Back to top of this page.....</a></DL></A></DL></A></OL>

SANS - Studies of Artificial Neural Systems</a>

Santa Fe Institute</a>

Science at the Edge of Chaos</a>

Sciences of Complexity</a>

Self-Organizing Emergent Behavior</a>

Self-Organizing Systems</a>

Selfreplicating shar archives</a>

Simple Classifier System</a>

Simulation Resources, Inc. (SRI)

a neural network simulator

written at the University of Stuttgart, Germany

Society for Nonlinear Dynamics and Econometrics</a> (Rutgers)

Society for the study of artificial intelligence and simulation of behavior (AISB)</a>

Softbots</a> (University of Washington)

Software Agents Mailing List</a> by thread

Software Complexity in Rule-Based Systems</a>

starlogo</a> A simple complex systems simulations implemented in logo

Stochastic Complexity</a>

Studies in Nonlinear Dynamics and Econometrics</a>

a genetic algorithm package

Sutton's Reinforcement Learning Archive</a>

The Swarm Project</a> headed by Chris Langton.

Back to top of this page.....
Appendix A

<DT><DD><LI><A HREF="http://www.cs.brandeis.edu/~zippy/alife-library.html">Virtual ALife Library</A>
<DT><DD><LI><A HREF="http://life.csu.edu.au/VI_complex/library1.html">Virtual Library on Complex Systems</A>
<DT><DD><LI><A HREF="http://web-bou.iapc.net/~koops/vivarium/vivarium.html">Vivarium</A> is an interactive simulation of the evolution of behavior
<DT><DD><LI><A HREF="http://www.cs.brandeis.edu/dept/index.html">Volen National Center for Complex Systems</A> at Brandeis University
</DL>
</DL>
</DD>
</OL>
</DD>
</DT><DD><DL><DL><A HREF="#aa">Back to top of this page.....</A></DL>

</DL>

</DD>
</DT><DD><HL><HR>
</HL>

</DL>

</DD>
</OL>

<DT><DD><LI><A HREF="http://fuzine.mt.cs.cmu.edu/mlm/signidr94.html">Web Agent Related Research</A> (Carnegie Mellon University)
<DT><DD><LI><A HREF="http://wissgl.weizmann.ac.il/physics/phys_nlin.html">Weizmann Institute (Israel) non-linear dynamics</A>
<DT><DD><LI><A HREF="http://www.nd.com/welcome/whatisnn.htm">What is an Artificial Neural Network?</A>
<DT><DD><LI><A HREF="http://www.nonlin.tu-muenchen.de/chaos/ww.html">Who Is Who Handbook of Nonlinear Dynamics</A> (Munchen)
<DT><DD><LI><A HREF="http://newciv.org/worldtrans/whole.html">Whole Systems</A>
<DT><DD><LI><A HREF="http://www.seattleanticoch.edu/wholeSystem/">Whole Systems Design</A>
<DT><DD><LI><A HREF="http://netq.rowland.org/sw/swhp.html">Stewart Wilson's Homepage</A>
<DT><DD><LI><A HREF="ftp://ftp.Germany.EU.net/pub/research/softcomp/Alife/packages/winlife">WinLife</A> A nice implementation of John Conway's "Life Game"
<DT><DD><LI><A HREF="http://www.ai.univie.ac.at/oefai/nn/servers.html">WWW Neural Network Home Pages</A> (University of Vienna)
</DL>
</DL>
</DD>
</DT><DD><DL><DL><A HREF="#aa">Back to top of this page.....</A></DL>
</DL>
</DD>
</DT><DD><HL><HR>
</HL>
</DL>

</DD>
</DT><DD><DL><DL><A NAME="X">X</A></DL>
</DL>
</DD>
</DT><DD><HL><HR>
</HL>
</DL>

</DD>
</DT><DD><OL>
<DT><DD><LI><A HREF="http://ai.toronto.edu/pub/xerion">Xerion</A> a neural network simulator
<DT><DD><LI><A HREF="ftp://parcftp.xerox.com/pub/dynamics">Xerox PARC</A> some papers on the evolution of cooperative behaviour
<DT><DD><LI><A HREF="http://www.ccsf.caltech.edu/ismap/image.html">Xmorphia</A> morphogenesis from a reaction-diffusion system
</OL>

177
Appendix A

<DT><DD><DL><LI><A HREF="http://penguin.phy.bnl.gov/www/xtoys">Xtoys</A> programs for X windows that self-organized criticality
<DT><DD><DL><LI><A HREF="http://www.mindspring.com/~zsol/nintro.html">ZSolutions</A> an introduction to neural networking
</DL></DD></DL>
<DL><DT><DD><DL><LI><A HREF="#aa">Back to top of this page</A>.....</DL>
</DL>
<DL><HR></DL>
<DL><H1><A NAME="Y">Y</A></H1>
</DL>
<DL><OL>
<DL><DT><DD><DL><LI><A HREF="#aa">Back to top of this page</A>.....</DL>
</DL>
<DL><HR></DL>
<DL><H1><A NAME="Z">Z</A></H1>
</DL>
<DL><DT><DD><DL><LI><a href="http://alife.santafe.edu:80/~joke/zooland/">Zooland</a>
<DT><DD><DL><LI><A HREF="http://www.d.umn.edu/~cbusch/toolbox.html">ZooLife</A> an alife application written in C++ for UNIX
</DL>
<DL><HR></DL>
<DL><DT><DD><DL><LI><A HREF="#aa">Back to top of this page</A>.....</DL>
</DL>
<DL><HR></DL>
<HR size=7mm>
<p> Last Update: 18 March, 1996
<address>a href = "mailto:ilachina@cna.org" ilachina@cna.org</address>
</p>
Appendix B: Glossary of Terms

Adaptation

Any change in the structure or function of an entity (say, a biological organism) that allows it to survive and reproduce more effectively in its environment.

Algorithmic Complexity

A measure of the complexity of a problem. Typically defined as the size of the smallest program that computes the given problem or that generates a complete description of it.

Animats

Artificial animals consisting of both software and hardware. Typically designed to be able to adapt to their environment over time.

Artificial Life

This is not a concept that is yet ready to be rigorously defined. The most concise, but still far from rigorous definition, is simply: life as synthesized by man rather than by nature. One of the basic tenets of this still-infant field is the belief that life is not unique to its biological (and, as yet, only known) form, but is a more general property of the organization of matter. Artificial life explores life as it could be as opposed to life as we know it to be.

Attractor

Dissipative dynamical systems are characterized by the presence of some sort of internal "friction" that tends to contract phase-space volume elements. Contraction in phase space allows such systems to approach a subset of the phase-space called an attractor as the elapsed time grows large. Attractors therefore describe the long-term behavior of a dynamical system. Steady state (or equilibrium) behavior corresponds to fixed-point attractors, in which all trajectories starting from the appropriate basin-of-attraction eventually converge onto a single point. For linear dissipative dynamical systems, fixed point attractors are the only possible type of attractor. Nonlinear systems, on the other hand, harbor a much richer spectrum of attractor types. For example, in addition to fixed-points, there may exist periodic attractors such as limit cycles. There is also an intriguing class of chaotic attractors called strange attractors that have a complicated geometric structure (see Chaos and Fractals).

Autonomous (or Adaptive-) Agent

An entity that, by sensing and acting upon its environment, tries to fulfill a set of goals in a complex, dynamic environment
Appendix B

- it can sense the environment through its sensors and act on the environment through its actuators

- it has an internal information processing and decision making capability

- it can anticipate future states and possibilities, based on internal models (which are often incomplete and/or incorrect)
  - this anticipatory ability often significantly alters the aggregate behavior of the system of which an agent is part

- an agent's goals can take on diverse forms:
  - desired local states
  - desired end goals
  - selective rewards to be maximized
  - internal needs (or motivations) that need to be kept within desired bounds

- since a major component of an agent's environment consists of other agents, agents spend a great deal of their time adapting to the adaptation patterns of other agents

Autoplectic Systems

Consider a dynamical system whose behavior appears random or chaotic. There are two ways in which an apparent randomness can occur: (1) external noise, so that if the evolution of the system is unstable, external perturbations amplify exponentially with time — such systems are called homoplectic; (2) internal mechanisms, so that the randomness is generated purely by the dynamics itself and does not depend on any external sources or require that randomness be present in the initial conditions — such systems are called autoplectic systems. An example of an autoplectic system is the one-dimensional, two-state, two neighbor Cellular Automaton rule-30, starting from a single non-zero site. The temporal sequence of binary values starting from that single non-zero initial seed are completely random, despite the fact that the evolution is strictly deterministic and the initial state is ordered.

Autopoiesis

Autopoiesis literally means "self-reproduction," and expresses a fundamental complementarity between structure and function. More precisely, the term refers to the dynamics of non-equilibrium structures; that is, organized states (sometimes also called dissipative structures) that remain stable for long periods of time despite matter and energy continually flowing through them. A vivid example of a nonequilibrium structure is the Great Red Spot on Jupiter, which is essentially a gigantic whirlpool of gases in Jupiter's upper
atmosphere. This vortex has persisted for a much longer time (on the order of centuries) than the average amount of time any one gas molecule has spent within it.

**Backpropagation Algorithm**

The backpropagation algorithm is a learning rule for multi-layered Neural Networks, credited to Rumelhart and McClelland. The algorithm gives a prescription for adjusting the initially randomized set of synaptic weights (existing between all pairs of neurons in each successive layer of the network) so as to maximize the difference between the network's output of each input fact and the output with which the given input is known (or desired) to be associated. The backpropagation rule takes its name from the way in which the calculated error at the output layer is propagated backwards from the output layer to the N\textsuperscript{th} hidden layer to the (N-1)\textsuperscript{th} hidden layer, and so on. Because this learning process requires us to to "know" the correct pairing of input-output facts beforehand, this type of weight adjustment is called supervised learning.

**Basin of Attraction**

The basin of attraction is the ensemble of points P such that if the trajectory starts from P it approaches the Attractor.

**Bifurcation**

The splitting into two modes of behavior of a system that previously displayed only one mode. This splitting occurs as a control parameter is continuously varied. In the Logistic Equation, for example, a period-doubling bifurcation occurs whenever all the points of period-2\textsuperscript{n} cycle simultaneously become unstable and the system becomes attracted to a new period-2\textsuperscript{n+1} cycle.

**Boolean Function**

A function that maps an n-tuple of binary values – (x\textsubscript{1}, x\textsubscript{2}, ..., x\textsubscript{n}) where xi = 0 or 1 for all i – to another binary value (either 0 or 1). There are clearly 2\textsuperscript{2}\textsuperscript{n} possible Boolean functions that can defined for a given n-tuple.

**Cantor Set**

A simple example of a Fractal set of points having noninteger Hausdorff Dimension. For example, the triadic Cantor set is constructed as follows: take the unit interval (= [0,1]) and generate a new set by deleting the open interval (1/2, 2/3); that is, by deleting the middle third. Generate a new set by deleting the middle thirds (1/9, 2/9) and (7/9, 8/9) from the previous set with the middle third removed. The Cantor set is essentially what remains of the unit interval in the limit of generating successive "middle third" deleted sets from the original set. It can be shown that the Fractal Dimension of this set is approximately equal to 0.6309.
Appendix B

Catastrophe Theory

Catastrophe theory, introduced by Thom in the 1960s, is a mathematical formalism for modeling nonlinear systems whose behavior is determined by the actions of a small number of driving parameters. In particular, it applies to systems that undergo either gradual or sudden changes in behavior due to gradually changing forces. It has been applied to many problems in mathematics, physics and the social sciences. Thom called the sudden changes that take place in a system "catastrophes" and developed a theory as a method of analyzing and classifying these changes. Thom's theorem asserts that the stationary state behavior of all systems that have up to four control parameters (or input variables) and two behavior (or output) variables, and which also have an associated potential function, can be described using one of seven elementary catastrophes.

Cellular Automata

Cellular automata (CA) are a class of spatially and temporally discrete, deterministic mathematical systems characterized by local interaction and an inherently parallel form of evolution. First introduced by von Neumann in the early 1950s to act as simple models of biological self-reproduction, CA are prototypical models for complex systems and processes consisting of a large number of identical, simple, locally interacting components. The study of these systems has generated great interest over the years because of their ability to generate a rich spectrum of very complex patterns of behavior out of sets of relatively simple underlying rules. Moreover, they appear to capture many essential features of complex self-organizing cooperative behavior observed in real systems. Although much of the theoretical work with CA has been confined to mathematics and computer science, there have been numerous applications to physics, biology, chemistry, biochemistry, and geology, among other disciplines. Some specific examples of phenomena that have been modeled by CA include fluid and chemical turbulence, plant growth and the dendritic growth of crystals, ecological theory, DNA evolution, the propagation of infectious diseases, urban social dynamics, forest fires, and patterns of electrical activity in neural networks. CA have also been used as discrete versions of partial differential equations in one or more spatial variables.

Cellular Games

A cellular game is a dynamical system in which sites of a discrete lattice play a "game" with neighboring sites. Strategies may be deterministic or stochastic. Success is usually judged according to a universal and fixed criterion. Successful strategies persist and spread throughout the lattice; unsuccessful strategies disappear.

Chaos

Deterministic chaos refers to irregular or chaotic motion that is generated by nonlinear systems evolving according to dynamical laws that uniquely determine the state of the system at all times from a knowledge of the system's previous history. It is important to point out that the chaotic behavior is due neither to external sources of noise nor to an infinite number of degrees-of-freedom nor to quantum-mechanical-like uncertainty. Instead, the
source of irregularity is the exponential divergence of initially close trajectories in a bounded region of phase-space. This sensitivity to initial conditions is sometimes popularly referred to as the "butterfly effect," alluding to the idea that chaotic weather patterns can be altered by a butterfly flapping its wings. A practical implication of chaos is that its presence makes it essentially impossible to make any long-term predictions about the behavior of a dynamical system: while one can in practice only fix the initial conditions of a system to a finite accuracy, their errors increase exponentially fast.

**Chaotic Control**

It has recently been suggested that the extreme sensitivity of chaotic systems to small perturbations to initial conditions (the so-called "butterfly effect") can be exploited to stabilize regular dynamic behaviors and to effective "direct" chaotic trajectories to a desired state. This is a capability that has no counterpart in nonchaotic systems for the ironic reason that the trajectories in nonchaotic systems are stable and thus relatively impervious to desired control. A recent survey article (Grebogi, Ott, et. al.) lists applications for communications (in which chaotic fluctuations can be put to use to send controlled, pre-planned signals), for physiology (controlling chaos in heart rhythms), for fluid mechanics and chemical reactions. As another recent example, a few years ago NASA used small amounts of residual hydrazine fuel to steer the ISEE-3/ICE spacecraft to its rendezvous with a comet 50 million miles away. This was possible because of the sensitivity of the three-body problem of celestial mechanics to small perturbations.

**Classifier Systems**

Classifier systems were introduced by John Holland as an attempt to apply Genetic Algorithms to cognitive tasks. They are similar to production systems of the "if...then" variety in artificial intelligence. A classifier system typically consists of (1) a set of detectors (or input devices) that provide information to the system about the state of the external environment, (2) a set of effectors (or output devices) that transmit the classifier's conclusions to the external environment, (3) a set of rules (or classifiers), consisting of a condition and action, and (4) a list of messages. Learning is supervised as in multilayered Neural Networks.

**Class-P Problems**

The Computational Complexity of a problem is defined as the time it takes for the fastest program running on a universal computer – as measured in number of computing steps, say N – to compute the solution to the problem. The complexity is then classified according to how fast N grows as a function of the problem size, s. The first non-trivial class of problems – class-P – consists of problems for which the computation time increases as some polynomial function of s. Problems that can be solved with polynomial-time algorithms are called tractable; if they are solvable but are not in the class-P, they are called intractable.

**Co-Adaptation/Co-Evolution**

The evolutionary process of a biological species in nature is often described as though that species were trying to adapt to a fixed environment. However,
such a description only crudely approximates what really happens. In nature, the "environment" consists of both a relatively (but not completely) stable physical environment as well as other species of organisms that are simultaneously trying to adapt to their environment. The actions of each of these other species typically affects the actions of all other species that occupy the same physical environment. In biology (and hence Artificial Life and studies involving Genetic Algorithms), the terms "co-adaptation" and "co-evolution" are sometimes used to refer to the fact that all species simultaneously co-adapt and co-evolve in a given physical environment.

**Complex Adaptive Systems**

Macroscopic collections of simple (and typically nonlinearly) interacting units that are endowed with the ability to evolve and adapt to a changing environment.

**Complexity**

An extremely difficult "I know it when I see it" concept to define, largely because it requires a quantification of what is more of a qualitative measure. Intuitively, complexity is usually greatest in systems whose components are arranged in some intricate difficult-to-understand pattern or, in the case of a dynamical system, when the outcome of some process is difficult to predict from its initial state. In its lowest precisely when a system is either highly regular, with many redundant and/or repeating patterns or when a system is completely disordered. While over 30 measures of complexity have been proposed in the research literature, they all fall into two general classes:

- **Static Complexity** – which addresses the question of how an object or system is put together (i.e. only purely structural informational aspects of an object), and is independent of the processes by which information is encoded and decoded

- **Dynamic Complexity** – which addresses the question of how much dynamical or computational effort is required to describe the information content of an object or state of a system

Note that while a system’s static complexity certainly influences its dynamical complexity, the two measures are not equivalent. A system may be structurally rather simple (i.e. have a low static complexity), but have a complex dynamical behavior.

**Computational Complexity**

Computational complexity measures the time and memory resources that a computer requires in order to solve a problem. A somewhat more robust measure may be defined by invoking the Universal Turing Machine. The Computational Complexity of a problem is then defined as the time it takes for the fastest program running on a universal computer (as measured in number of computing steps) to compute the solution to the problem.
Appendix B

Computational Irreducibility

Much of theoretical physics has traditionally been concerned with trying to find "shortcuts" to nature. That is to say, with trying to find methods that are able to reproduce a final state of a system by knowing the initial state but without having to meticulously trace out each step from the initial to final states. The fact that we can write down a simple parabola as a path a thrown object makes in a gravitational field is an example of an instance where this might be possible. Clearly such shortcuts ought to be possible in principle if the calculation is more sophisticated than the computations the physical system itself is able to make. But consider a computer. Because a computer is itself physical system, it can determine the outcome of its evolution only by explicitly following it through. No shortcut is possible. Such computational irreducibility occurs whenever a physical system can act as a computer. In such cases, no general predictive ability is possible. Computational irreducibility implies that there is a highest level at which abstract models of physical systems can be made. Above that level, one can model only by explicit simulation.

Computational Universality

Computational universality is a property of a certain class of computers such that changes in input alone allow any computable function to be evaluated without any change in internal construction. Universal computers can thus simulate the operation of any other computer, given that their input is suitably coded. Conway's Life Game, for example, has been shown to be a universal computer. This means that with a proper selection of initial conditions (i.e. the initial distribution of "live" and "dead" cells), Life can be turned into a general purpose computer. This fact fundamentally limits the overall predictability of Life's behavior. The Halting Theorem, for example, asserts that there cannot exist a general algorithm for predicting when a computer will halt its execution of a given program. Given that Life is a universal computer – so that the Halting theorem applies – this means that one cannot, in general, predict whether a particular starting configuration of live and dead cells will eventually die out. No shortcuts are possible, even in principle.

Conservative Dynamical Systems

In contrast to Dissipative Dynamical Systems, conservative systems preserve Phase Space volumes and hence cannot display any attracting regions in phase space; there can be no fixed points, no limit cycles and no strange attractors. There can nonetheless be chaotic motion in the sense that points along particular trajectories may show sensitivity to initial conditions. A familiar example of a conservative system from classical mechanics is that of a Hamiltonian system.

Cost Function

In optimization problems, the cost function measures how good a particular solution to the problem is; the higher its value the better the solution. Also called the fitness function.
Appendix B

Coupled-Map Lattices

Generic Cellular Automata (CA) are dynamical systems in which space, time and the local state space are all discretized. Coupled-map lattices are simple generalizations of CA in which space and time remain discrete, but in which the individual site values are allowed to take on continuous values.

Criticality

"Criticality" is a concept borrowed from thermodynamics. Thermodynamic systems generally get more ordered as the temperature is lowered, with more and more structure emerging as cohesion wins over thermal motion. Thermodynamic systems can exist in a variety of phases – gas, liquid, solid, crystal, plasma, etc. – and are said to be critical if poised at a phase transition. Many phase transitions have a critical point associated with them, that separates one or more phases. As a thermodynamic system approaches a critical point, large structural fluctuations appear despite the fact the system is driven only by local interactions. The disappearance of a characteristic length scale in a system at its critical point, induced by these structural fluctuations, is a characteristic feature of thermodynamic critical phenomena and is universal in the sense that it is independent of the details of the system's dynamics. (See Self-Organized Criticality)

Crossover Operator

One of three basic genetic operations used in Genetic Algorithms. Reproduction makes a set of identical copies of a given chromosome, where the number of copies depends on the chromosome's fitness. The crossover operator exchanges subparts of two chromosomes, where the position of the crossover is randomly selected, and is thus a crude facsimile of biological sexual recombination between two single-chromosome organisms. The mutation operator randomly flips one or more bits in the chromosome, where the bit positions are randomly chosen.

Dissipative Structure

An organized state of a physical system whose integrity is maintained while the system is far from equilibrium. Example: the great Red-Spot on Jupiter. Dissipative Dynamical Systems Dissipative systems are dynamical systems that are characterized by some sort of "internal friction" that tends to contract phase space volume elements. Phase space contraction, in turn, allows such systems to approach a subset of the space called an Attractor (consisting of a fixed point, a periodic cycle, or Strange Attractor), as time goes to infinity.

Edge-of-Chaos

The phrase "edge-of-chaos" refers to the idea that many complex adaptive systems, including life itself, seem to naturally evolve towards a regime that is delicately poised between order and chaos. More precisely, it has been used as a metaphor to suggest a fundamental equivalence between the dynamics of phase transitions and the dynamics of information processing. Water, for example, exists in three phases: solid, liquid and gas. Phase-transitions denote the boundaries between one phase and another. Universal computation – that is, the ability to perform general purpose computations and which is arguably
an integral property of life - exists between order and chaos. If the behavior of a system is too ordered, there is not enough variability or novelty to carry on an interesting calculation; if, on the other hand, the behavior of a system is too disordered, there is too much noise to sustain any calculation. Similarly, in the context of evolving natural ecologies, "edge-of-chaos" refers to how - in order to successfully adapt - evolving species should be neither too methodical nor too whimsical or carefree in their adaptive behaviors. The best exploratory strategy of an evolutionary "space" appears at a phase transition between order and disorder. Despite the intuitive appeal of the basic metaphor, note that there is currently some controversy over the veracity of this idea.

Emergence

Emergence refers to the appearance of higher-level properties and behaviors of a system that - while obviously originating from the collective dynamics of that system's components - are neither to be found in nor are directly deducible from the lower-level properties of that system. Emergent properties are properties of the "whole" that are not possessed by any of the individual parts making up that whole. Individual line of computer code, for example, cannot calculate a spreadsheet; an air molecule is not a tornado; and a neuron is not conscious. Emergent behaviors are typically novel and unanticipated.

Entropy

A measure of the degree of randomness or disorder in a system. Determines a system's capacity to evolve irreversibly in time. Specific definitions vary depending on the type of system considered. Examples: (1) in statistical systems, the entropy is proportional to the logarithm of the total number of possible states with the same energy as the state under consideration.; (2) in classical thermodynamics, the differential change in entropy of a system near equilibrium is the differential change in absorbed heat divided by the system temperature; (3) in nonlinear deterministic dynamical systems, the Kolmogorov-Sanai entropy is often used as a measure. It is defined as the sum of the positive Lyapunov Exponents of the system.

Ergodic System

An ergodic dynamical system is one whose trajectory eventually "covers" the entire phase space. Put another way, given any point P in the phase space, the trajectory will approach P arbitrarily closely for sufficiently large times t.

Ergodic Theory

A branch of applied mathematics that uses statistical concepts to describe average properties of deterministic dynamical systems. The ergodic hypothesis (which asserts that a phase-space average of a measurable X is equal to its time-average) provides the basis for classical statistical mechanics. Attempts at providing a rigorous mathematical proof of the ergodic hypothesis include Poincare's recurrence theorem (which asserts that a trajectory will return to any neighborhood of its initial state if one waits long enough) and the ergodic theorems of Birkhoff and von Neumann.
Appendix B

Evolution

A general term referring to the dynamical unfolding of behavior over time. Darwinian evolution refers to the unfolding of higher (i.e. more complex) life forms out of lower life forms.

Evolutionary Programming

Evolutionary programming is essentially an application of genetic algorithms to computer programs. Typically the genome is represented by a LISP expression, so that what evolves is a population of programs, rather than bit-strings as in the case of a usual genetic algorithm. For references see Koza [179] and the WWW sources listed in appendix A.

Evolutionary Stable Strategy

A concept from a generalized form of Game Theory. Animals are endowed with a finite set of possible strategies that they can use in their interactions with other animals. Strategies may be "pure," in which the animal acts according to a prescribed set of instructions in all contexts, or "mixed," in which the animal adopts different strategies with different probabilities. The evolutionary stable strategy (ESS) is a strategy, or set of strategies such that if it is adopted by all animals no other strategy can invade the population.

Finite Automata

Human languages can, conceptually, be regarded as a set of rules for constructing sequences of symbols according to a fixed set of rules of composition in order to convey meaning. One can therefore consider using a Cellular Automaton as a formalism for studying the abstract properties of language. To be more precise, a finite automaton $M$ is defined to consist of a finite alphabet $A$, a finite set of states $X$, and a state-transition function $f: X \times A \rightarrow X$ that gives the next state given the current state and the current input symbol. (There is also a set $T$ in $X$, which is the set of final or accepting states of the automaton.)

Fitness Landscape

A name for the landscape representing the fitness measure (or Cost Function) of a problem. Examples: Traveling Salesman Problem, survivability of a real or virtual creature.

Flicker- (or 1/ƒ-) Noise

Whenever the power spectral density, $S(\tilde{f})$, scales as $\tilde{f}^\alpha$, the system is said to exhibit $1/\tilde{f}$-noise (or flicker-noise). Despite being found almost everywhere in nature – $1/\tilde{f}$-noise has been observed in the current fluctuations in a resistor, in highway traffic patterns, in the price fluctuations on the stock exchange, in fluctuations in the water level of rivers, to name just a few instances – there is currently no fundamental theory that adequately explains why this same kind of noise should appear in so many diverse kinds of systems. What is clear is that since the underlying dynamical processes of these systems are so different, the common bond cannot be dynamical in nature, but can only be a kind of "logical dynamics" describing how a system's degrees-of-freedom all
interact. self-organized criticality may be a fundamental link between temporal scale invariant phenomena and phenomena exhibiting a spatial scale invariance. Bak, et. al., argue that $1/f$ noise is actually not noise at all, but is instead a manifestation of the intrinsic dynamics of Self-Organized Critical systems.

**Fractals**

Fractals are geometric objects characterized by some form of self-similarity; that is, parts of a fractal, when magnified to an appropriate scale, appear similar to the whole. Coastlines of islands and continents and terrain features are approximate fractals. The Strange Attractors of nonlinear dynamical systems that exhibit deterministic Chaos typically are fractals.

**Fractal Dimension**

Suppose a set can be covered by a finite number $N$ of segments of length $L$. There is a simple scaling relationship between these two numbers. For a line segment, $L$ grows as $1/R$; for a square, $L$ grows as $1/R^2$; for a cube, $L$ grows as $1/R^3$, and so on. The fractal dimension $D$ is defined by generalizing this intuitive scaling: $D = \lim_{r \to 0} \ln(N)/(\ln(1/R))$, where $\ln(x)$ is the natural logarithm. Sometimes also called the Hausdorff dimension or the Kolmogorov capacity.

**Frustration**

In Spin Glasses, a phenomenon in which individual magnetic moments receive competing ordering instructions via different routes, because of the variation of the interaction between pairs of atomic moments with separation.

**Fuzzy Logic**

Fuzzy set theory provides a formalism in which the conventional binary logic based on choices "yes" and "no" is replaced with a continuum of possibilities that effectively embody the alternative "maybe". Formally, the characteristic function of set $X$ defined by $f(x) = 1$ for all $x$ in $X$ and $f(x) = 0$ for all $x$ not in $X$ is replaced by the membership function $0 < m(x) < 1$ for all $x$ in $X$. The mathematics of fuzzy set theory was originated by L. A. Zadeh in 1965.

**Genetic Algorithms**

Genetic algorithms are a class of heuristic search methods and computational models of adaptation and evolution based on natural selection. In nature, the search for beneficial adaptations to a continually changing environment (i.e. evolution) is fostered by the cumulative evolutionary knowledge that each species possesses of its forebears. This knowledge, which is encoded in the chromosomes of each member of a species, is passed from one generation to the next by a mating process in which the chromosomes of "parents" produce "offspring" chromosomes. Genetic algorithms mimic and exploit the genetic dynamics underlying natural evolution to search for optimal solutions of general combinatorial optimization problems. They have been applied to the travelling salesman problem, VLSI circuit layout, gas pipeline control, the parametric design of aircraft, neural net architecture, models of international security, and strategy formulation.
Appendix B

Genotype

The genetic instruction code of an individual.

Hamiltonian System

A dynamical system that conserves volumes in phase space. Examples include mechanical oscillators without friction and the motion of a planet.

Hausdorff Dimension

For an operational definition of Hausdorff dimension, proceed as follows: Suppose a set can be covered by a finite number \( N \) of segments of length \( L \). There is a simple scaling relationship between these two numbers. For a line segment, \( L \) grows as \( 1/R \); for a square, \( L \) grows as \( 1/R^2 \); for a cube, \( L \) grows as \( 1/R^3 \), and so on. The Hausdorff dimension \( D \) is defined by generalizing this intuitive scaling: \( D = \lim_{R \to 0} \ln(N)/\ln(1/R) \), where \( \ln(x) \) is the natural logarithm. Sometimes also called the fractal dimension or the Kolmogorov capacity.

Hierarchy

Hierarchies consist of levels each of which include all lower levels; i.e. systems within systems within systems...within the total system in question. Evolution in complex systems leads to differentiation in multilevel hierarchic systems.

Homoclinic Point

A point in Phase Space of a nonlinear dynamical system that evolves to a point of unstable equilibrium in infinite time. Homoclinic Orbit The ensemble of points in the Phase Space of a nonlinear dynamical system that all evolve to a point of unstable equilibrium after an infinite time.

Hopf-Bifurcation

In the Logistic Map, a fixed point may lose its stability by splitting (or bifurcating) into a pair of points that form a period two orbit. Another common way in which a point may become unstable is by effectively turning into a small circle that then increases in size, deforms and becomes unstable as the controlling parameter is increased. This is called the Hopf Bifurcation.

Hypercycle

A scenario for the origin of self-replicating molecular systems proposed by Manfred Eigen. The scenario involves template-instructed replicating cycles consisting of feedback loops in which molecule A generates molecule B, molecule B generates molecule C, and molecule C generates molecule A, and so on.

Information Dimension

Partition a \( d \)-dimensional Phase Space into boxes of volume \( \varepsilon \). The probability of finding a point of an Attractor in box \( i \) is \( p_i(\varepsilon) = N_i(\varepsilon)/N(\varepsilon) \), where \( N_i(\varepsilon) \) is the number of points in the \( i \)th box and \( N(\varepsilon) \) is the total
number of non-empty boxes. \( p_i(\varepsilon) \) is the relative frequency with which the \( i \)th box is visited. The amount of information required to specify the state of the system to within an accuracy "\( \varepsilon \)" (or, equivalently, the information gain in making a measurement that is uncertain by an amount "\( \varepsilon \)"), is given by. The information dimension, \( D_0 \), of an attractor is then defined to be
\[
D_0 = \lim_{\varepsilon \to 0} \frac{I(\varepsilon)}{\ln(1/\varepsilon)}.
\]

Information Theory

Like the physically primitive notions of mass and energy of a particle, the information content, \( I \), of an arbitrary measurement or message composed of particular symbol sequence, is itself a primitive concept. While the roots of information theory extend back to the definition of the classical entropy of a physical system introduced by Clausius in 1864 and Boltzman's probabilistic reinterpretation of classical entropy in 1896, the mathematical formalism for measuring \( I \) is due largely to a seminal 1948 paper by Claude E. Shannon. Within the context of sending and receiving messages in a communication system, Shannon was interested in finding a measure of the information content of a received message. Shannon's approach was to obtain a measure of the reduction of uncertainty given some a-priori knowledge of the symbols being sent. Suppose we are given \( N \) different and a-priori equally likely possible outcomes. A measure of the information gain, \( I \), is obtained by required that \( I \) be additive for independent events. That is to say, if there are two independent sets of outcomes \( N_1 \) nd \( N_2 \), so that the total number of outcomes is \( N = N_1 \times N_2 \), it is required that \( I(N_1 \times N_2) = I(N_1) + I(N_2) \). This requirement is uniquely satisfied by the function \( I = c \log(N) \), where "\( c \)" is an arbitrary constant.

Intermittency

A term used in the study of nonlinear dynamical systems describing the changes between quiet, regular periods of activity (called the laminar phase) and periods of wild, chaotic oscillation (called bursts). Intermittency is a common route to chaos in physical systems.

Kolmogorov Entropy

The Kolmogorov entropy (or K-entropy) is a useful measure by which to characterize chaotic motion in an arbitrary-dimensional phase space. Loosely speaking, the K-entropy is proportional to the rate at which information about the state of a dynamical system is lost in the course of time. It is related to the average Lyapunov Exponent, which measures the exponential rate of divergence of nearby trajectories.

Lattice Gas Models

Lattice gases are micro-level rule-based simulations of macro-level fluid behavior. The Navier-Stokes Equations, the fundamental equations describing incompressible fluid flow, are in general analytically intractable. Lattice-gas models provide a powerful new tool in modeling real fluid behavior. The idea is to reproduce the desired macroscopic behavior of a fluid by modeling the underlying microscopic dynamics. In order to achieve an Emergence of a suitable macrodynamics out of a discrete microscopic substrate, one must have three basic ingredients: (1) local thermodynamic equilibrium, (2)
conservation laws, and (3) a "scale separation" between the levels at which the microscopic dynamics takes place (among kinetic variables living on a micro-lattice) and the collective motion itself appears (defined by hydrodynamical variable on a macro-lattice). Another critical feature is the symmetry of the underlying lattice. While there are many basic variants of the model, one can show that there is a well-defined minimal set of rules that define a lattice-gas system whose macroscopic behavior reproduces that predicted by the Navier-Stokes equations exactly.

**Life Game**

Invented by the mathematician John Conway, Life is arguably the most widely known Cellular Automaton rule. It was extensively popularized by Martin Gardner in his "Mathematical Games" department in Scientific American in the early 1970s. Life is "played" using the eight nearest-neighbors on a lattice, and consists of (1) seeding the lattice with some pattern of "live" and "dead" cells, and (2) simultaneously (and repeatedly) applying the following three rules to each cell of the lattice at discrete time steps:

- **Birth**: replace a previously dead cell with a live one if exactly 3 of its neighbors are alive

- **Death**: replace a previously live cell with a dead one if either (1) the living cell has no more than one live neighbor (i.e. it dies of isolation), or (2) the living cell has more than three neighbors (i.e. it dies of overcrowding)

- **Survival**: retain living cells if they have either 2 or 3 neighbors

One of the most intriguing patterns in Life is an oscillatory propagating pattern known as the "glider." It consists of 5 "live" cells and reproduces itself in a diagonally displaced position once every four iterations. When the states of Life are projected onto a screen in quick succession by a fast computer, the glider gives the appearance of "walking" across the screen. The propagation of this pseudo-stable structure can also be seen as a self-organized emergent property of the system.

**Limit-Cycle**

An Attractor describing regular (i.e. periodic or quasi-periodic) temporal behavior.

**Lindenmeyer (or L) Systems**

L-systems were introduced by Aristid Lindenmeyer in 1968 as a model for the cellular development of filamentous plants. In simplest terms, L-systems consist of production rules for rewriting abstract strings of symbols. They can be thought of as generalized Cellular Automata in which the number of sites can increase over time.

**Logistic Equation**

The logistic map is one of the simplest (continuous and differentiable) nonlinear systems that captures most of the key mechanisms responsible for
producing deterministic chaos. It is a one-dimensional nonlinear discrete difference equation with a single control parameter, \( x_{n+1} = ax_n(1-x_n) \), where \( 0 < x_0 < 1 \) and \( 0 < a < 4 \). The logistic equation undergoes a sequence of period-doubling bifurcations followed by regions of deterministic chaos as \( a \) is varied between the values 0 and 4. Some aspects of this behavior - such as the ratio of bifurcation intervals as chaos is approached - are Universal; that is, are independent of the details of the system.

**Lotka-Volterra Equations**

In 1926, Volterra proposed a simple model for the predation of one species by another to explain the oscillatory level of certain fish in the Atlantic. If \( N(t) \) is the prey population and \( P(t) \) is the predator population at time \( t \) then Volterra's model is \( dN/dt = N(a-bP) \), \( dP/dt = P(cN-d) \), where \( a, b, c, \) and \( d \) are positive constants. The model assumes: (1) prey in absence of predation grows linearly with \( N \) (i.e. in Malthusian fashion); (2) predation reduces prey's growth rate by a term proportional to the prey and predation populations; (3) the predator's death rate, in the absence of prey, decays exponentially; (4) the prey's contribution to the predator's growth rate is proportional to the available prey as well as to the size of the predator population. The system of equations is known as the Lotka-Volterra equations because Lotka derived the same equations in 1920 for a chemical reaction he believed to exhibit periodic behavior.

**Lyapunov Exponent**

A fundamental property of chaotic dynamics is sensitivity to small changes in initial conditions. Initially closely separated starting conditions evolving along regular dynamical trajectories diverge only linearly in time; a chaotic evolution, on the other hand, leads to exponential divergence in time. Lyapunov exponents quantify this divergence by measuring the mean rate of exponential divergence of initially neighboring trajectories. A trajectory of a system with a negative Lyapunov exponent is stable and will converge to an attractor exponentially with time. The magnitude of the Lyapunov exponent determines how fast the attractor is approached. A trajectory of a system with a positive Lyapunov exponent is unstable and will not converge to an attractor. The magnitude of the positive Lyapunov exponent determines the rate of exponential divergence of the trajectory.

**Markov Process**

A Markov process is a process for which, if the present is given, the future and past are independent of each other. More precisely, if \( t_1 < ... < t_n \) are parameter values, and if \( 1 < j < n \), then the sets of random variables \( \{ x(t_1), ..., x(t_{j-1}) \} \) and \( \{ x(t_{j-1}), ..., x(t_n) \} \) are mutually independent for given \( x(t_j) \). Equivalently, the conditional probability distribution of \( x(t_n) \) for given \( x(t_1), ..., x(t_{j-1}) \) depends only on the specified value of \( x(t_{j-1}) \) and is in fact the conditional probability distribution of \( x(t_n) \), given \( x(t_{j-1}) \). An important and simple example is the Markov chain, in which the number of states is finite or denumerably infinite.
Appendix B

**Maximum Entropy**

The principle of maximum entropy states that when one has only partial information about the probabilities of possible outcomes of an experiment, one should choose the probabilities so as to maximize the uncertainty about the missing information. Put another way, since entropy is a measure of randomness, one should choose the most random distribution subject to whatever constraints are imposed on the problem.

**Mean-Field Theory**

In a mean field approximation a system is assumed to be determined by the average properties of the system as a whole. In a mean-field-theoretic description of a thermodynamic system, for example, all particles are considered to contribute equally to the potential at each site. Therefore, the mean field theory essentially assumes the intermolecular interaction to be of infinite range at all temperatures. The mean field theories are qualitatively quite successful in that they predict the existence of critical points and power law dependence of the various thermodynamic quantities near the critical point. They generally become more quantitatively successful as the dimensionality of the system increases.

**Multifractal**

The simplest fractal sets are characterized by some form of self-similarity, in which parts, when magnified by a constant \( r \), appear similar to the original whole. The more general class of fractals are really multi-scale fractals, or multifractals, which are characterized by multiple subdivisions of the original into \( N \) objects, each magnified by by a different factor \( r_i, i=1,2,...,N \).

**Navier-Stokes Equations**

These are a set of analytically intractable coupled nonlinear partial differential equations describing fluid flow.

**Neural Networks**

Neural nets represent a radical new approach to computational problem solving. The methodology they represent can be contrasted with the traditional approach to artificial intelligence (AI). Whereas the origins of AI lay in applying conventional serial processing techniques to high-level cognitive processing like concept-formation, semantics, symbolic processing, etc. – or in a top-down approach – neural nets are designed to take the opposite – or bottom-up – approach. The idea is to have a human-like reasoning emerge on the macro-scale. The approach itself is inspired by such basic skills of the human brain as its ability to continue functioning with noisy and/or incomplete information, its robustness or fault tolerance, its adaptability to changing environments by learning, etc. Neural nets attempt to mimic and exploit the parallel processing capability of the human brain in order to deal with precisely the kinds of problems that the human brain itself is well adapted for.
Nonlinearity

If $f$ is a nonlinear function or an operator, and $x$ is a system input (either a function or variable), then the effect of adding two inputs, $x_1$ and $x_2$, first and then operating on their sum is, in general, not equivalent to operating on two inputs separately and then adding the outputs together; i.e. the whole is not necessarily equal to the sum of its parts. Dissipative nonlinear dynamic systems are capable of exhibiting self-organization and chaos.

NP-Hard Problems

A class of problems, known as nondeterministic polynomial time – or class-NP – problems, that may not necessarily be solvable in polynomial time, but the actual solutions to which may be tested for correctness in polynomial time.

NP-Complete

Just as there are universal computers that, given a particular input, can simulate any other computer (see Universal Computer), there are NP-complete problems that, with the appropriate input, are effectively equivalent to any NP-hard problem of a given size. For example, Boolean "satisfiability" – i.e. the problem of determining truth values of the variable's of a Boolean expression so that the expression is true – is known to be an NP-complete problem.

Order Parameter

An order parameter is a scalar or vector parameter associated with a continuous phase transition that determines the physical nature of the transition. It has the value zero in the random state (typically above the transition temperature) and takes on a nonzero value in the ordered state (typically below the transition). In the case of a fluid, for example, the order parameter is a scalar and is the difference in density between the liquid and vapor phases.

Percolation Theory

Percolation theory represents the simplest model of a disordered system. Consider a square lattice, where each site is occupied randomly with probability $p$ or empty with probability $1-p$. Occupied and empty sites may stand for very different physical properties. For simplicity, let us assume that the occupied sites are electrical conductors, the empty sites represent insulators, and that electrical current can flow between nearest neighbor conductor sites. At low concentration $p$, the conductor sites are either isolated or form small clusters of nearest neighbor sites. Two conductor sites belong to the same cluster if they are connected by a path of nearest neighbor conductor sites, and a current can flow between them. At low $p$ values, the mixture is an insulator, since a conducting path connecting opposite edges of the lattice does not exist. At large $p$ values, on the other hand, many conduction paths between opposite edges exist, where electrical current can flow, and the mixture is a conductor. At some concentration in between, therefore, a threshold concentration $p_c$ must exist where for the first time electrical current can percolate from one edge to the other. Below $p_c$, we have an insulator; above $p_c$ we have a conductor. The threshold concentration is
called the percolation threshold, or, since it separates two different phases, the critical concentration.

**Petri Nets**

Petri nets are abstract models used to represent parallel systems and processes. They are typically described using directed graphs (i.e., graphs whose edges are depicted by arrows showing a direction of information flow). More precisely, a petri net is a seven-tuple \((P, T, V, f, g, N, m)\), where (1) \(P\) is a nonempty finite set of nodes, (2) \(T\) is a nonempty finite set of transitions, (3) \(V\) is a valuation space \([0,1]\), (4) \(f\) is a binary function used in determining the connections from nodes to transitions (i.e., \(f: P \times T \to V\), and if \(f(p,t)=1\) then node \(p\) connects to transition \(t\), otherwise not), (5) \(g\) is a binary function used in determining the transitions to connect to nodes (i.e., \(g: T \times P \to V\) and a connection is made from \(t\) to \(p\) if and only if \(g(t,p)=1\)), (6) \(N\) is a set of markings \([0,1,2,...]\), and (7) \(m\) is the initial marking function, \(m: P \to N\).

**Phase Space**

A mathematical space spanned by the dependent variables of a given dynamical system. If the system is described by an ordinary differential flow the entire phase history is given by a smooth curve in phase space. Each point on this curve represents a particular state of the system at a particular time. For closed systems, no such curve can cross itself. If a phase history a given system returns to its initial condition in phase space, then the system is periodic and it will cycle through this closed curve for all time. Example: a mechanical oscillator moving in one-dimension has a two-dimensional phase space spanned by the position and momentum variables.

**Phase Transition**

An abrupt change in a system's behavior. A common example is the gas-liquid phase transition undergone by water. In such a transition, a plot of density versus temperature shows a distinct discontinuity at the critical temperature marking the transition point. Similar behavior can be seen in systems described by ordinary differential flows and discrete mappings. In nonlinear dynamical systems, the transition from self-organizing to chaotic behavior is sometimes referred to as a phase transition (or, more specifically, as an order-disorder transition).

**Phenotype**

The overall attributes of an organism arising from the interaction of its Genotype with the environment.

**Poincare Map**

A dynamical system is usually defined as a continuous flow, that is (1) is completely defined at all times by the values of \(N\) variables \(- x_1(t), x_2(t), ..., x_N(t)\), where \(x_i(t)\) represents any physical quantity of interest, and (2) its temporal evolution is specified by an autonomous system of \(N\), possibly coupled, ordinary first-order differential equations. Once the initial state is specified, all future states are uniquely defined for all times \(t\). A convenient method for visualizing continuous trajectories is to construct an equivalent
Appendix B

discrete-time mapping by a periodic "stroboscopic" sampling of points along a trajectory. One way of accomplishing this is by the so-called Poincare map (or surface-of-section) method. Suppose the trajectories of the system are curves that live in a three-dimensional Phase Space. The method consists essentially of keeping track only of the intersections of this curve with a two-dimensional plane placed somewhere within the phase space.

Prisoner's Dilemma

The prisoner's dilemma is a two person non-zero-sum game that has been widely used in experimental and theoretical investigations of cooperative behavior. Two persons suspected of a crime are caught, but there is not enough evidence to sentence them unless one of them confesses. If they are both quiet (or cooperate, C), both will have to be released. If one confesses (defects, D) but the other does not, the one who confesses will be released but the other will be imprisoned for a long time. Finally, if both confess, both will be imprisoned, but for a shorter time. It is assumed that the prisoner's make their respective choices separately and independently of one another. If the game is "played" once, each player find defection to be the optimal behavior, regardless of what his opponent chooses to do. Finding the optimal strategy to follow over time, however, is considerably more difficult.

Probabilistic CA

Cellular Automata for which the deterministic state transitions are replaced with specifications of the probabilities of cell-value assignments. For such systems, the focus of analysis shifts from studying evolutions of arbitrary initial states to studying ensembles of trajectories.

Punctuated Equilibrium

A theory introduced in 1972 to account for what the fossil record appears to suggest are a series of irregularly spaced periods of chaotic and rapid evolutionary change in what are otherwise long periods of evolutionary stasis. Some Artificial Life studies suggest that this kind of behavior may be generic for evolutionary processes in complex adaptive systems.

Quasiperiodic

Characterizes behavior of a dynamical system that is almost, but not quite, periodic. Quasiperiodic regions of phase space are frequently linked together to form a Strange Attractor. The transition between such quasiperiodic regions is characterized by the crossing of a Homoclinic Point. Quasiperiodicity often results when nonlinear dynamical systems are driven by periodic driving forces with periods that are incommensurate with (i.e. not a rational fraction of) the system response time.

Random Boolean Networks

A size N random Boolean network of size k generalizes the basic binary Cellular Automata model by evolving each site variable $x_i = 0$ or 1 according to a randomly selected Boolean Function of k inputs. Since there are two choices for every combination of states of the k inputs at each site, the Boolean function is randomly selected from among the $2^k$ possible Boolean
functions of k inputs. This model was first introduced by Kauffman in 1969 in a study of cellular differentiation in a biological system (binary sites were interpreted as elements of an ensemble of genes switching on and off according to some set of random rules). Since its conception, however, related models have found wide application in an increasingly large domain of diverse problems. Such models of strongly disordered systems exhibit remarkable and unexpected order.

**Reaction-Diffusion Models**

Reaction-diffusion systems, the first studies of which date back to the 1950s, often exhibit a variety of interesting spatial patterns that evolve in self-organized fashion. One of the most famous reaction-diffusion systems – widely regarded as the prototypical example of oscillating chemical reactions – is the so-called Belousov-Zhabotinskii (or BZ) reaction. The BZ model involves the reaction of bromate ions with an organic substrate (typically malonic acid) in a sulfuric acid solution with cerium (or some other metal-ion catalyst). When this mixture is allowed to react exothermally at room temperature, interesting spatial and temporal oscillations (i.e. chemical waves) result. The system oscillates, changing from yellow to colorless to back to yellow about twice a minute with the oscillations typically lasting for over an hour (until the organic substrate is exhausted). A number of Cellular Automata models have been found that exhibit BZ-like spatial waves.

**Relativistic Information Theory**

Relativistic information theory is a concept introduced by Jumarie and has been suggested as a possible formalism for describing certain aspects of military command and control processes by Woodcock and Dockery. The basic idea is that a generalized entropy is endowed with four components, so that it is equivalent to a four-vector and may be transformed by a Lorentz transformation (As in relativity). These four components consists of: (1) the external entropy of the environment (H_e), (2) the internal entropy of the system (H_s), (3) system goals, and (4) the internal transformation potential, which measures the efficiency of the system's internal information transformation. An additional factor, called the organizability, plays the role of "velocity." Woodcock and Dockery show that it is possible to use relativistic information theory to compare the relative command and control system response of two command structures to the world around them. The quantity of interest is \( \frac{dH_s}{dH_e} \), or the rate of change of the internal information environment with respect to changes in the surrounding environment.

**Scaling Laws**

Theoretical studies of critical phenomena have focused on predicting the value of critical exponents. One of the most important ideas is the scaling hypothesis. This hypothesis is model-independent and applicable to all critical systems. The underlying assumption is that the long-range correlation of the order parameter, such as the density fluctuation in a fluid system near the critical temperature, is responsible for all singular behavior. This assumption leads to a particular functional form for the equation of state near the critical point.
Appendix B

Search Space

The variation of the Cost Function can be imagined to be a landscape of potential solutions to a problem where the height of each feature represents its cost. This landscape is sometimes referred to as the search space.

Self-Organization

The spontaneous emergence of macroscopic nonequilibrium organized structure due to the collective interactions among a large assemblage of simple microscopic objects.

Self-Organized Criticality

Self-organized criticality (SOC) describes a large body of both phenomenological and theoretical work having to do with a particular class of time-scale invariant and spatial-scale invariant phenomena. Fundamentally, SOC embodies the idea that dynamical systems with many degrees of freedom naturally self-organize into a critical state in which the same events that brought that critical state into being can occur in all sizes, with the sizes being distributed according to a power-law. The kinds of structures SOC seeks to describe the underlying mechanisms for look like equilibrium systems near critical points (see Criticality) but are not near equilibrium; instead, they continue interacting with their environment, "tuning themselves" to a point at which critical-like behavior appears. Introduced in 1988, SOC is arguably the only existing holistic mathematical theory of self-organization in complex systems, describing the behavior of many real systems in physics, biology and economics. It is also a universal theory in that it predicts that the global properties of complex systems are independent of the microscopic details of their structure, and is therefore consistent with the "the whole is greater than the sum of its parts" approach to complex systems. Put in the simplest possible terms, SOC asserts that complexity is criticality. That is to say, that SOC is nature's way of driving everything towards a state of maximum complexity.

Simulated Annealing

A mathematical technique for general combinatorial optimization problems. The name comes from the physical process of annealing, during which a material is first heated and then slowly cooled. During annealing, the component atoms of a material are allowed to settle into a lower energy state so that a more stable arrangement of atoms is maintained throughout the cooling process.

Solitons

A mathematically appealing model of real particles is that of solitons. It is known that in a dispersive medium, a general wave form changes its shape as it moves. In a nonlinear system, however, shape-preserving solitary waves exist.

Spatio-Temporal Chaos

A large class of spatially extended systems undergoes a sequence of transitions leading to dynamical regimes displaying chaos in both space and time. In the same way as temporal chaos is characterized by the coexistence of a large
number of interacting time scales, spatio-temporal chaos is characterized by having a large number of interacting space scales. Examples of systems leading to spatio-temporal chaos include the Navier-Stokes Equations and reaction-diffusion equations. Coupled-map Lattices have been used for study.

**Spin Glasses**

A magnetic material whose magnetic magnets respond to both ferromagnetic and antiferromagnetic interactions causing frustration, so that not all the constraints necessary to minimize the system's overall energy can be simultaneously satisfied. There are exponentially stable states, but finding the global ground state is an NP-Hard optimization problem.

**Strange Attractors**

Describes a form of long-term behavior in dissipative dynamical systems. A strange attractor is an attractor (see Attractor) that displays sensitivity to initial conditions. That is to say, an attractor such that initially close points become exponentially separated in time. This has the important consequence that while the behavior for each initial point may be accurately followed for short times, prediction of long time behavior of trajectories lying on strange attractors becomes effectively impossible. Strange attractors also frequently exhibit a self-similar or fractal structure.

**Symbolic Dynamics**

Symbolic dynamics is a tool that is used to obtain a coarse-grained representation of dynamical orbits consisting of discrete-symbol sequences. This is done by first partitioning the phase space into a finite number of cells $C_1, C_2, ..., C_N$ and and focusing on the successive cell-to-cell transitions of the trajectory. The states of the cells, $S(C_1), S(C_2), ..., S(C_N)$, are treated as symbols of an N-letter alphabet. Looked at in this way, the continuous dynamics thus induces on the partition a symbolic dynamics describing how the letters of the alphabet evolve in time.

**Synergetics**

Synergetics refers to what can loosely be called the "European" (vice US) approach to the study of complex systems. Consider a complex system (that is, a system composed of many individual parts) that is controlled from the outside in some manner by a control parameter (say, the system is driven by a constant influx of energy and/or matter). As the control parameter is changed, the system's state can become unstable and be replaced by a new state characterized by particular kinds of spatial, temporal or functional structures. Synergetics consists of strategies of describing what happens when the macroscopic state of systems undergoes a qualitative change. More colloquially, "synergy" is used to refer to how the action of two or more entities ("parts") can achieve an effect that cannot be achieved by any of the parts alone (see Emergence).

**Topological Dimension**

The topological dimension of object X is an integer defining the number of coordinates needed to specify a given point of X. A single point therefore a
topological dimension equal to zero; a curve has dimension one, a surface has dimension two, and so on.

**Universality**

Universal behavior, when used to describe the behavior of a dynamic system, refers to behavior that is independent of the details of the system's dynamics. It is a term borrowed from thermodynamics. According to thermodynamics and statistical mechanics the critical exponents describing the divergence of certain physical measurables (such as specific heat, magnetization, or correlation length) are universal at a phase transition in that they are essentially independent of the physical substance undergoing the phase transition and depend only on a few fundamental parameters (such as the dimension of the space).

**Unstable Equilibrium**

A stationary state of a dynamical system such that an arbitrarily small perturbation can cause a disturbance of arbitrarily large magnitude. Example: an egg poised on the vertex of a cone.
Appendix C: Recommended Reading

Listed below are some recommended introductory texts and popularizations of nonlinear dynamics and complex systems theory:


- *Complexity: Metaphors, Models and Reality*, edited by G. A. Cowan, D. Pines and D. Meltzer, Addison-Wesley, 1994. This is a collection of short, basic research papers by practitioners of complex systems theory. Each paper is followed by excerpts of comments made during a panel discussion. Although the papers are generally presented at a technical level, the collection provides an excellent overview of complex systems theory.

discussion of complex systems theory at the Santa Fe Institute, set against the canyonlands and history of northern New Mexico.


References


11. J. Arquilla and D. Ronfeldt, "Cyberwar is coming!", Comparative Strategy, Volume 12, 1993, 141.


References


25. J. C. Bezdek and S. K. Pal, editors, Fuzzy Models for Pattern Recognition: Methods that


References


References


References


210
References


References


References


References


References


References

173. S. Kauffman and W. Mcready, "Technological evolution and adaptive organizations,"


References


221. M. Mitchell, P. T. Hraber and J. P. Crutchfield, "Revisiting the edge of chaos: evolving cellular automata to perform


235. Operations, Department of the Army FM 100-5, June 1993.


References


References


References


313. Satosi Watanabe, Knowing and Guessing: A Quantitative Study of Inference and Information, John Wiley and Sons.


228


