John H. Holland's general theories of adaptive processes apply across biological, cognitive, social, and computational systems.

BY STEPHANIE FORREST AND MELANIE MITCHELL

Adaptive Computation:
The Multidisciplinary Legacy of John H. Holland

IN AUGUST 2015, Professor John H. Holland passed away in Ann Arbor, MI, where he had served on the University of Michigan faculty for more than 50 years. John, as he was known universally to his colleagues and students, leaves behind a long legacy of intellectual achievements.

As a descendant of the cybernetics era, he was influenced by the work of John von Neumann, Norbert Wiener, W. Ross Ashby, and Alan Turing, all of whom viewed computation as a broad, interdisciplinary enterprise. Holland thus became an early proponent of interdisciplinary approaches to computer science and an active evangelist of what is now called computational thinking, reaching out enthusiastically to psychologists, economists, physicists, linguists, philosophers, and pretty much anyone he came in contact with. As a result, even though he received what was arguably one of the world's first computer science Ph.D. degrees in 1959,[23] his contributions are

sometimes better known outside computer science than within.

Holland is best known for his invention of genetic algorithms (GAs), a family of search and optimization methods inspired by biological evolution. Since their invention in the 1960s, GAs have inspired many related methods and led to the thriving field of evolutionary computation, with widespread scientific and commercial applications. Although the mechanisms and applications of GAs are well known, they were only one offshoot of Holland's broader motivation—to develop a general theory of adaptation in complex systems.

Here, we consider this larger framework, sketching the recurring themes that were central to Holland's theory of adaptive systems: discovery and dynamics in adaptive search; internal models and prediction; exploratory modeling; and universal properties of complex adaptive systems.

Discovery and Dynamics in Adaptive Search

Holland's goal of developing a general theory of adaptation was spurred by both his early work on computer models of Hebbian learning[25] and his reading of Ronald Fisher's classic book, *The Genetical Theory of Natural Selection*, which integrated genetics with Darwinian selection.[8] As he read further in evolutionary biology, economics, game theory, and control theory, Holland recognized that adaptation is central to

>> key insights

- Adaptation is an open-ended dynamic process involving populations of agents operating in perpetually novel environments, continually changing their behavior to improve performance and enhance their chance of survival.

- Adaptive systems create, update, and use internal models of their environment to make predictions; successful models create valid homomorphisms of their environment.

- Exploratory models of complex adaptive systems lead to insights about essential mechanisms and principles, a use of modeling that is different from simply making accurate predictions.

all of these fields; they all concern populations of agents that must continually obtain information from uncertain, changing environments and use it to improve performance and enhance the chance of survival.

Holland also recognized that adaptation must be continual and open ended. In his view, adaptive systems never achieve a state of equilibrium or final "optimum" configuration. This emphasis on open-ended, non-equilibrium dynamics was in stark contrast with the mainstream approach (at the time) in all these fields—the belief that solving for stable equilibrium dynamics was the scientific goal. Holland's contrary view was that a system in stable equilibrium is essentially dead.

Underlying Holland's theory of adaptation are the following core ideas:

Populations, sampling, and implicit parallelism. Evolution is a form of search that leverages statistics to direct population dynamics. Initially, the population is an independent sample of many individuals (from the space of all possible individuals) and over time the sample is increasingly biased toward the high-fitness regions of the search space. In addition, the population can be viewed as an implicit sample of the much larger space of traits exhibited by those individuals. Holland termed this implicit large-scale sampling of traits "implicit parallelism." Evolutionary dynamics biases these samples over time toward high-fitness regions of the search space.

Building blocks and recombination. In a population undergoing adaptation, individuals can be decomposed into building blocks—sets of traits that are the evolutionary "atoms" of an individual's fitness or performance. (As an example from biology, Holland often mentioned the Krebs cycle, a core cellular metabolic pathway that is highly conserved across living systems.) Successful individuals are discovered in stages, first by finding useful building blocks through stochastic sampling, and over time recombining them to create higher-fitness individuals out of

more complex building blocks.

Exploitation vs. exploration. Successful adaptation requires maintaining a balance between exploitation, in which successful building blocks propagate in a population, and exploration, in which existing building blocks are recombined or mutated in new ways.

Inspired by Bellman[3] and others, Holland formalized the exploitation-vs.-exploration trade-off as an idealized "two-armed bandit" problem. Given a slot machine with two arms, each of which has an unknown payoff probability, how should you allocate trials (pulls) between the arms so as to maximize your total payoff? Holland argued that the optimal strategy allocates trials to the observed best arm at

a slightly higher rate than an exponential function of the trials allocated to the observed worse arm.[11,12]

Further, Holland showed that in populations undergoing adaptation, building blocks are analogous to arms on a multi-armed bandit. Evaluating an individual in an environment is analogous to pulling selected arms on a multi-armed bandit, where the arms correspond to each of the building blocks making up that individual.

The question of balancing exploitation and exploration—how to optimally allocate trials to different arms based on their observed payoffs—now becomes the question of how to sample optimally in the vast space of possible building blocks, based on their

estimated contribution to fitness. Evolution deals in populations of individuals, of course, not building blocks. There is no explicit mechanism that keeps statistics on how building blocks contribute to fitness. Holland's central idea here is that nearly optimal building-block sampling occurs implicitly, as an emergent property of evolutionary population dynamics.

Holland's early papers[10,11] and his influential 1975 book *Adaptation in Natural and Artificial Systems*[12] developed a general, formal setting in which these ideas could be expressed mathematically. It was this formalization that led to the invention of genetic algorithms that featured stochastic population-based search, as well as crossover between individuals as a critical operation that allows successful building blocks to be recombined and tested in new contexts.

However, Holland's aim was more general than a new class of algorithms. He aspired to develop an interdisciplinary theory of adaptation, one that would inform, say, biology as much as computer science.[5] The later, successful application of genetic algorithms to real-world optimization and learning tasks was, for Holland, just icing on the cake. His broader view of adaptation has inspired many and engendered criticism from others, leading to spirited intellectual debates and controversies. Most controversial is the extent to which it can be demonstrated, either empirically or mathematically, that the behavior of adaptive systems is actually governed by Holland's proposed principles. Regardless of one's position on this question, the ideas are compelling and provide a set of unifying concepts for thinking about adaptation.

Internal Models and Prediction

Internal models are central to Holland's theory of adaptive systems. He posited that all adaptive systems create and use internal models to prosper in the face of "perpetual novelty."

Models can be tacit and learned over evolutionary time, as in the case of bacteria swimming up a chemical gradient, or explicit and learned over a single lifespan, as in the case of cognitive systems that incorporate experience into internal representations through learning. In Holland's view, the key activity of an adaptive agent involves building and refining these data-driven models of the environment.

In his second book, *Induction*,[14] Holland and his co-authors proposed inductive methods by which cognitive agents can construct—and profit from—internal models by combining environmental inputs and rewards with stored knowledge. In their framework, an internal model defines a set of equivalence classes over states of the world, together with a set of transition rules between the equivalence classes, all of which are learned over time based on environmental rewards (or punishments). The many-to-one mapping from states of the world to states of the model (the equivalence classes) is called a homomorphism. Models that form valid homomorphisms with the part of the world being modeled allow the system to make accurate predictions. In Holland's conception, the equivalence classes are initially very general, as with, say, "moving object" and "stationary object." Through experience and learning, these classes can be specialized into more useful and precise subclasses, as in, say, "insect" and "nest." Over time, the adaptive system builds up a default hierarchy of rules covering general cases and refinements for specific classes.

At the time of Holland's work, the idea of default hierarchies was prevalent in knowledge representation systems. Holland made two key contributions to this literature. The first was his emphasis on homomorphisms as a formal way to evaluate model validity, an idea that dates back to W. Ross Ashby's *An Introduction to Cybernetics*.[2] Holland's student Bernard Ziegler developed this idea into a formal theory of computer modeling and simulation.[30] Even today, these early homomorphic theories of modeling stand as the most elegant approach we know of to characterize how consistent a model is with its environment and how an intelligent agent, human or artificial, can update a model to better reflect reality.

Holland's second key contribution was describing a computational mechanism, the "learning classifier system,"[13,21] to illustrate how a cognitive system could iteratively build up a detailed and hierarchical model of its environment to enhance survival. The key learning elements of this method, the "bucket-brigade" algorithm, combined with a genetic algorithm, presaged many of the ideas in modern reinforcement learning, notably the temporal-difference methods introduced by Sutton and Barto.[28]

Holland's inspiration for classifier systems came from several different disciplines, including Hebbian learning, artificial intelligence, evolutionary biology, economics, psychology, and control theory. Knowledge representation in the form of a population of "if-then" rules was a natural choice due to its popularity in AI at the time, as well as Holland's early work on modeling Hebbian cell assemblies: "In Hebb's view, a cell assembly makes a simple statement: If such and such an event occurs, then I will fire for a while at a high rate."[29] The if-then rules, when activated, compete to post their results on a shared "message list," serving as the system's short-term memory, again inspired by Hebb's work, as well as by AI blackboard systems of the day. Unlike blackboard systems, however, new rules are generated automatically in a trial-and-error fashion, and can be selected and recombined by a genetic algorithm.

Successful rules are strengthened over time if their predictions lead to positive rewards from the environment (and weakened otherwise) through a credit-assignment method called the bucket-brigade algorithm, in which rules gaining rewards from the environment or from other rules, transfer some of their gains to earlier-firing "stage-setting" rules that set up the conditions for the eventual reward. Holland credited Arthur Samuel's pioneering work on machine learning applied to checkers[26] as a key inspiration for these ideas.

Holland was primarily interested in how the two learning mechanisms—discovery of new rules and apportioning credit to existing rules—could work together to create useful default hierarchies of rules. He emphasized that the competition inherent in the learning and action mechanisms would allow the system to adapt to a continually changing environment without losing what it had learned in the past.

Holland put it this way: "Competition among rules provides the system with a graceful way of handling perpetual novelty. When a system has strong rules that respond to a particular situation, it is the equivalent of saying that it has certain well-validated hypotheses … New rules do not interfere with the system's action in well-practiced situations but wait gracefully in the wings as hypotheses about what to do under novel circumstances."[15]

Although Holland proposed classifier systems as an executable theory of inductive processes in cognition, other researchers took them further, developing applications to areas as diverse as poker playing,[27] control of gas pipeline transmission,[9] and modeling the stock market.[24] (See Booker et al.[4] for more on practical applications of classifier systems.) Today, other reinforcement learning methods are more popular for real-world decision and control problems, but classifier systems can perhaps be viewed as an early stage-setting method that foreshadowed these later approaches.

Exploratory Modeling

Given that Holland believed the ability to learn and manipulate internal models was essential for any adaptive system, it is no surprise that he viewed modeling as essential for scientific inquiry.

Today, we use computational models both for prediction—by analyzing data via statistical models—and for understanding how systems work—by probing the effects of hypothesized underlying mechanisms. This latter use of models for enhancing understanding was dear to Holland's heart. In his view, the key to science is to understand the mechanisms that cause a system to behave in a certain way, an aspiration that goes well beyond data-fitting methods, which typically focus on the aggregate behavior of a system.

For example, a purely statistical model that describes the boom-and-bust pattern of the stock market does not address the underlying mechanisms that lead to these cycles through the collective actions of myriad individual buy/sell decisions. In contrast, the genetic algorithms for which Holland is so famous are exploratory models of mechanism; they provide

> Holland is best known for his invention of genetic algorithms, a family of search and optimization methods inspired by biological evolution.

a simple computational framework in which to explore the dynamics of Darwinian evolution and whether the basic mechanisms of variation, differential reproduction, and heredity are sufficient to account for the richness of the natural world.

Holland's work focused on exploratory models—those that explore basic principles and mechanisms, even if they do not make specific or detailed predictions.[16] Such models can show generically how certain behaviors could be produced. Holland pioneered a style of modeling that has come to be known as "individual-based" or "agent-based," in which every component of a system is represented explicitly (such as every trader in a stock market system or every cell in an immune system model) and has a dynamic internal state. In such models, each agent has its own behavior rules it can modify over time, through learning. In order to capture the behavior of systems under spatial constraints, these models are often defined over spatial structures (such as networks or simple grids).

The exploratory models championed by Holland were not intended to provide detailed, domain-specific predictions. They were meant instead to explore general mechanisms of complex systems and thus provide insights that might lead to more specific, detailed models. Such idealized models are akin to what philosopher Daniel Dennett has called "intuition pumps."[6]

The emphasis on exploratory models to build intuitions was an important theme of Holland's work, and he often quoted Sir Arthur Eddington's remark on the occasion of the first experimental test of Albert Einstein's theory of relativity: "The contemplation in natural science of a wider domain than the actual leads to a far better understanding of the actual."[7] This view of modeling is not typical. For many scientists, models are useful only to the extent they generate accurate predictions, a perspective that rules out the very kind of exploratory modeling that interested Holland the most.

Universal Properties of Complex Adaptive Systems

Holland was interested in a broad array of adaptive systems, including immune systems, ecologies, financial

markets, cities, and the brain. In the early 1980s, he teamed up with a small group of scientists, primarily physicists, with a sprinkling of economists and biologists, to discuss what properties this wide swath of "complex adaptive systems" have in common. The discussions helped define the intellectual mission of the Santa Fe Institute, the first institution dedicated to developing a science of complexity, as well as the other complexity institutes that followed. Holland brought to these discussions his lifelong study of adaptation and a reminder that serious theories about complexity would need to look deeper than phenomenological descriptions but also account for the "how" and "why" of these systems.

As the discussions about complex adaptive systems matured, a consensus developed about their basic properties. Such systems are composed of many components with nonlinear interactions; are characterized by complex emergent behavior; exhibit higher-order patterns; operate at multiple (and often nested) spatial and temporal scales, with some behavior conserved across all scales and other behaviors changing at different scales; and are adaptive, with behavioral rules continually adjusted through evolution and learning. Although this list is far from a formal characterization of complex adaptive systems, most people working in the field today are interested in systems that have these properties.

In the early 1990s, Holland teamed up with other Santa Fe Institute researchers, including several economists, to tackle the mismatch between the predictions of rational expectations theory—the dominant theory in economics at the time—and empirically observed stock-market behaviors. In brief, most economic theory of the day assumed that all participants in an economy or financial market are fully rational, acting to maximize their individual gain. In real life, however, actors in economies and markets are rarely wholly rational, and financial markets often deviate from rationality, as in, say, speculative bubbles and crashes.

The Santa Fe Institute Artificial Stock Market project[1,24] developed an exploratory model in which rational traders were replaced by adaptive traders—those who learn to forecast stock

Holland's view was that a system in stable equilibrium is essentially dead.

prices over time. The model tested for the emergence of three different trading strategies: fundamental, technical, and uninformed. The simulated market with adaptive trading agents was run many times, and the dynamics of price and trading volumes were compared to observed patterns in real markets. Holland and his collaborators found that the model's dynamics replicated several otherwise puzzling features of real-life markets.

Although the Santa Fe Institute Stock Market model was highly simplified, it was very influential and led to many follow-on projects. It demonstrated clearly the essential role that adaptation plays in complex systems and illustrated how Holland's theories of continual learning in response to intermittent feedback from the environment could be integrated into domain-specific settings.

Echo[17,22] was an even more ambitious exploratory model Holland and his collaborators developed during the 1990s. Echo formalized Holland's idealization of complex adaptive systems into a runnable computational system where agents evolved external markers (called tags) and internal preferences, then used them to acquire resources and form higher-level aggregate structures (such as trading relationships, symbiotic groups, trophic cascades, and interdependent organizations). Echo agents sometimes discovered mimicry and used it to deceive competitors. These complex interacting patterns arose throughout the model's execution, which started with a single minimal agent. When the runs stabilized, the resulting population of agents was found to reproduce several well-known patterns observed in nature, most famously the rank-frequency distribution of species diversity known as the Preston curve in ecology. In lay terms, Echo evolved populations where "most species are rare."

The broad scope of the model, together with its ability to produce easily identifiable and well-known patterns from nature, was appealing to immunologists, economists, and evolutionary biologists alike. Many of the insights behind the project are described in Holland's third book, *Hidden Order*.[16]

Holland's later books, *Emergence*,[18] *Signals and Boundaries*,[19] and *Complex-*

ity: A Very Short Introduction[20] show how the theories of adaptation Holland developed during the earlier part of his career fit into the larger landscape of complex systems research. In these works, he especially emphasized the open-ended nature of complex systems, where change and adaptation are continual, systems co-evolve with each other, and ecological niches arise and decay spontaneously depending on resource availability and competition.

Holland's focus on understanding the mechanisms by which complex patterns emerge and change, rather than simply characterizing the patterns themselves (such as describing chaotic attractors or power laws), reflected his determination to get to the heart of complex adaptive systems. This determination represents the best of science. Holland's willingness to tackle the most difficult questions, develop his own formalisms, and use mathematics and simulation to provide insight sets a high bar for scientists in all disciplines.

Conclusion

John Holland was unusual in his ability to absorb the essence of other disciplines, articulate grand overarching principles, and then back them up with computational mechanisms and mathematics. Unlike most researchers, Holland moved seamlessly among these three modes of thinking, developing models that were years ahead of their time. A close reading of his work reveals the antecedents of many ideas prevalent in machine learning today (such as reinforcement learning in non-Markovian environments and active learning). His seminal genetic algorithm spawned the field of evolutionary computation, and his insights and wisdom helped define what are today referred to as the "sciences of complexity."

Holland's many books and papers have influenced scientists around the world and across many disciplines. Closer to home, he introduced several generations of students at the University of Michigan to computation in natural systems, an idea that even today remains somewhat controversial, despite successes in genetic algorithms for engineering design, biomimicry for robotics, abstractions of pheromone trails in ant colonies for optimization methods, and mechanisms from immunology to improve computer security.

Behind the ideas is the man himself. Everyone who knew John personally remembers the gleam in his eye when encountering a new idea; his willingness to talk to anyone, no matter how famous or obscure; and his incredible generosity and patience. His personality and humanity were somehow inextricably entangled with his intellectual contributions. Since his death in 2015, many of Holland's former students and colleagues have movingly described their desire to emulate his personal qualities as much as his scientific excellence. His ideas, intellectual passion, and personal approach will serve as beacons for research in intelligent and complex systems for many years to come.

Acknowledgments

References

1. Arthur, W.B. Holland, J.H., LeBaron, B.D., Palmer, R.G., and Tayler, P. Asset pricing under endogenous expectations in an artificial stock market. In *The Economy As an Evolving Complex System II*, W.B. Arthur, S. Durlauf, and D. Lane, Eds. Addison-Wesley, Reading, MA, 1997.
2. Ashby, W.R. *An Introduction to Cybernetics.* Chapman & Hall, London, 1956.
3. Bellman, R.E. *Adaptive Control Processes: A Guided Tour.* Princeton University Press, Princeton, NJ, 1961.
4. Booker, L.B., Goldberg, D.E., and Holland, J.H. Classifier systems and genetic algorithms. *Artificial Intelligence 40*, 1–3 (1989), 235–282.
5. Christiansen, F.B. and Feldman, M.W. Algorithms, genetics, and populations: The schemata theorem revisited. *Complexity 3*, 3 (1998), 57–64.
6. Dennett, D.C. Elbow Room: The Varieties of Free Will Worth Wanting. MIT Press, Cambridge, MA, 1984.
7. Eddington, A. *The Nature of the Physical World: Gifford Lectures, 1927.* Cambridge University Press, Cambridge, U.K., 1927.
8. Fisher, R.A. *The Genetical Theory of Natural Selection: A Complete Variorum Edition.* Oxford University Press, Oxford, U.K.,1930.
9. Goldberg, D.E. *Computer-Aided Gas Pipeline Operation Using Genetic Algorithms and Rule Learning.* Ph.D. Dissertation, University of Michigan, Ann Arbor, MI, 1983; http://www.worldcat.org/title/computer-aided-gas-pipeline-operation-using-genetic-algorithms-and-rule-learning/oclc/70390568
10. Holland, J.H. Outline for a logical theory of adaptive systems. *Journal of the ACM 9*, 3 (1962), 297–314.
11. Holland, J.H. Genetic algorithms and the optimal allocation of trials. *SIAM Journal on Computing 2*, 2 (1973), 88–105.
12. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* University of Michigan Press, Ann Arbor, MI, 1975.
13. Holland, J.H. Escaping brittleness: The possibilities of general-purpose learning algorithms applied to parallel rule-based systems. In *Machine Learning: An Artificial Intelligence Approach, Volume 2*, R.S. Michalski, J.G. Carbonell, and T.M. Mitchell, Eds. Morgan-Kaufman, San Francisco, CA, 1986, 593–623.
14. Holland, J.H. *Induction: Processes of Inference, Learning, and Discovery.* MIT Press, Cambridge, MA, 1989.
15. Holland, J.H. Genetic algorithms. *Scientific American 267* (July 1992), 44–50.
16. Holland, J.H. *Hidden Order: How Adaptation Builds Complexity.* Perseus Books, New York, 1995.
17. Holland, J.H. Echoing emergence: Objectives, rough definitions, and speculations for Echo-class models. In *Complexity: Metaphor, Models, and Reality*, G.A. Cowen, D. Pines, and D. Meltzer, Eds. Perseus Books, New York, 1999, 309–342.
18. Holland, J.H. *Emergence: From Chaos to Order.* Oxford University Press, Oxford, U.K., 2000.
19. Holland, J.H. *Signals and Boundaries: Building Blocks for Complex Adaptive Systems.* MIT Press, Cambridge, MA, 2012.
20. Holland, J.H. *Complexity: A Very Short Introduction.* Oxford University Press, Oxford, U.K., 2014.
21. Holland, J.H. and Reitman, J.S. Cognitive systems based on adaptive algorithms. *ACM SIGART Bulletin 63* (June 1977), 49.
22. Hraber, P.T., Jones, T., and Forrest, S. The ecology of Echo. *Artificial Life 3*, 3 (1997), 165–190.
23. London, R.L. *Who Earned First Computer Science Ph.D.?* blog@cacm (Jan. 15, 2013); http://cacm.acm.org/blogs/blog-cacm/159591-who-earned-first-computer-science-phd/fulltext
24. Palmer, R.G., Arthur, W.B., Holland, J.H., LeBaron, B., and Tayler, P. Artificial economic life: A simple model of a stock market. *Physica D: Nonlinear Phenomena 75*, 1–3 (1994), 264–274.
25. Rochester, N., Holland, J.H., Haibt, L.H., and Duda, W. Tests on a cell assembly theory of the action of the brain, using a large digital computer. *IRE Transactions on Information Theory 2*, 3 (1956), 80–93.
26. Samuel, A.L. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development 3*, 3 (1959), 210–229.
27. Smith, S.F. *A Learning System Based on Genetic Adaptive Algorithms.* Ph.D. Dissertation, University of Pittsburgh, Pittsburgh, PA, 1980.
28. Sutton, A.G. and Barto, R.S. Time derivative models of Pavlovian reinforcement. In *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, M. Gabriel and J. Moore, Eds. MIT Press, Cambridge, MA, 1990, 497–537.
29. Waldrop, M.M. *Complexity: The Emerging Science at the Edge of Order and Chaos.* Simon and Schuster, New York, 1993.
30. Ziegler, B. *Theory of Modeling and Simulation.* Wiley Interscience, New York, 1976.

Stephanie Forrest (forrest@cs.unm.edu) is Distinguished Professor of Computer Science at the University of New Mexico, Albuquerque, NM, External Professor at the Santa Fe Institute, Santa Fe, NM, and a Fellow of the IEEE; John Holland was her Ph.D. advisor at the University of Michigan.

Melanie Mitchell (mm@pdx.edu) is Professor of Computer Science at Portland State University and External Professor at the Santa Fe Institute, Santa Fe, NM; John Holland was co-advisor for her Ph.D. at the University of Michigan.