

A Profoundly Repeated Pattern

The most philosophically compelling story of the 20th century, in a century overfilled with scientific achievement, must nonetheless be the “discovery” and exploitation of information. The information revolution that ensued not only touched the lives of every human on the planet, but also fundamentally revolutionized our philosophy of biology.

The person most directly at the center of this revolution in thought was Claude Elwood Shannon, an electrical engineer and computer scientist whose theories laid the groundwork for the electronic communications networks that now ring the earth. Claude Shannon died on February 25, 2001, after a long fight with Alzheimer’s disease. He was 84.

Most ecologists will recognize the name because of the Shannon-Wiener index, but many may not know of the pivotal role that he has played in the information revolution characterizing our age. His name and his ideas are, however, exceptionally well known to every young engineering student.

Because I am a practicing design engineer, as well as someone who dabbles in evolutionary ecology, allow me to write a bit about Shannon, not so much about the man, but of his ideas and the profound, transforming effect that they have had on every aspect of biology.

I have never believed in the indispensability of individuals, particularly of scientists. The truth is there to be discovered, and like all truths, it is patient. There is almost an irrelevancy to who actually gets there first. Intellectual climate and culture are undoubtedly more important than the names of the individuals involved. If Einstein or Newton or Darwin or Shannon had never lived, we would still know pretty much everything we know now. Nevertheless, Shannon’s work ranks with that of Einstein, Newton, and Darwin, and will be increasingly remembered in the same way.

Half a century ago, Shannon wrote two papers that remain monuments

in the fields of computer science and information theory. Shannon’s 1940 master’s thesis, “*A symbolic analysis of relay and switching circuits*,” used Boolean algebra to establish the first theoretical underpinnings of digital circuits and networks and, as a consequence, it has frequently been described as the most influential master’s thesis ever written. But to greatly exceed that work, Shannon rather single-handedly invented information theory during the earliest part of his tenure at Bell Labs. The profundity of this second effort simply cannot be exaggerated.

There are two definitions of information, and they should not be confused. The first is the more common: “information is that quality that encodes behavior.” No better illustration of this definition exists than Richard Lewontin’s (1974) diagram of the evolution of a lineage through a single generation as a series of transformational mappings (Fig. 1).

Nearly four billion years ago, code and behavior somehow spontaneously segregated themselves into two distinct informational and functional realms: nucleic acids and catalytic proteins, and in doing so, unleashed an enormous flexibility in informational design. Fifty-five years ago, with the construction of the first stored-program, von Neumann-architected computers, we recapitulated these same advantages by separating code from behavior in the digital machines that we now build.

This repetition of natural design in machinery, which has often been

done by people who had no knowledge that nature had earlier evolved the same mechanism, has proven to be an exceptionally common theme over the past century. The phenomenon has often allowed us to understand the natural process fully for the first time. It may be that we are required to first build a structure before we can truly understand its purpose or its value.

Information is certainly one of these attributes. Of the three qualities that characterize life—mass, energy, and information—it is information that defines life. Indeed, it can be argued that information does not exist in the absence of life (Eigen 1992).

In that regard, Richard Dawkins gets it exactly right when he writes:

A few years ago, if you had asked almost any biologist what was special about living things as opposed to nonliving things, he would have told you about a special substance called protoplasm. Protoplasm wasn’t like any other substance; it was vital, vibrant, throbbing, pulsating, “irritable” (a schoolmarmish way of saying responsive). . . . When I was a schoolboy, elderly textbook authors still wrote about protoplasm, although, by then, they really should have known better. Nowadays, you never hear or see the word. It is as dead as phlogiston and the universal aether. There is nothing special about the substances from which living things are made. Living

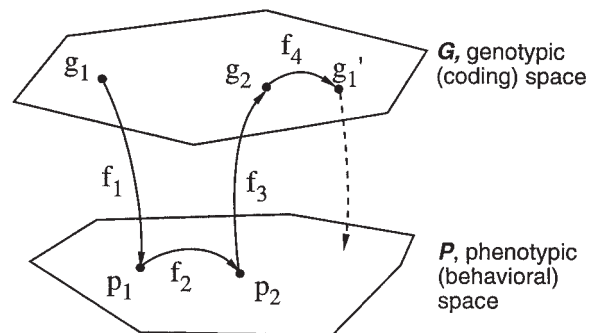


Fig. 1. Evolution of a lineage through a single generation, after Lewontin (1974). The various mappings are: f_1 is *developmental epigenesis*, f_2 is *selection*, f_3 is next generation *genetic representation*, and f_4 is *mutation*, in the largest possible sense.

things are collections of molecules, like everything else. What is special is that these molecules are put together in much more complicated patterns than the molecules of nonliving things, and this putting together is done by following programs, sets of instructions for how to develop, which the organisms carry around inside themselves. Maybe they do vibrate and throb and pulsate with “irritability,” and glow with “living” warmth, but these properties all emerge incidentally. What lies at the heart of every living thing is not a fire, not warm breath, not a “spark of life.” It is information, words, instructions. If you want a metaphor, don’t think of fires and sparks and breath. Think, instead, of a billion discrete, digital characters carved in tablets of crystal. If you want to understand life, don’t think about vibrant, throbbing gels and oozes, think about information technology.

(Dawkins 1987:111-112).

In 1948, Shannon published “A mathematical theory of communication” in the *Bell System Technical Journal*, giving birth to a new science, information theory. His motivation, however, was eminently practical: how to transmit messages while keeping them from becoming garbled by noise.

To analyze the problem properly, he realized that he first had to develop a precise definition of information, a dauntingly slippery concept. The information content of a message, Shannon argued, has nothing to do with its meaning, but rather with the number of 1’s and 0’s that it takes to transmit it.

This was a jarring notion to a generation of engineers who were accustomed to thinking of communication in terms of sending electromagnetic waveforms down a wire. “Nobody had come close to this idea before,” Robert Gallager, a long-time colleague, said in the *New York Times*’ obituary of Shannon. “This was not something somebody else would have

done for a very long time.” The overarching lesson was that the nature of the message did not matter—it could be numbers, words, music, video. Ultimately it was all just 1’s and 0’s (Johnson 2001).

Shannon defined information a second time, in a nonobvious way, as the metric:

$$I = -\log(p_i)$$

in direct imitation of Ludwig Boltzmann’s earlier definition of entropy

$$S = k \log W$$

where p_i is the probability of occurrence of the i th symbol in the symbol set, S is entropy, k is Boltzmann’s constant, and W the probability of a collection of system states.

By defining his information metric in this manner, what Shannon was explicitly defining was the *unexpectedness* of the i th symbol, or the level of *surprise* that accompanies its appearance.

Although it may seem odd on first hearing, Darwinian evolutionary biology, thermodynamics, and information theory are the same subject, and they have been philosophically intertwined from their very beginnings. The line of thought that leads from Darwin to Shannon is not only direct, but was highly intended by each of its participants.

Rudolph Clausius (1865) first defined the term entropy, S , as that fraction of energy that is lost to the pool of irrecoverable heat in every ordered transaction. Entropy literally means “in one turn.” The mental model that Clausius employed was that of the gearing system in a grist mill. Some fraction of the ordered energy in every turn of the primary drive shaft is never returned to the subsequent gears; rather, it is irrecoverably lost to an inaccessible pool of heat due to friction.

Ludwig Boltzmann (1872) almost immediately redefined Clausius’ entropy, not as a fluid-like, bulk quality of heat, as Carnot, Kelvin, Clausius, and Maxwell had envisioned the

process, but rather as a population of particles inexorably moving from a state of order to one of increasingly greater disorder. This is called the *microscopic interpretation* in physics, although it could just as easily be termed the *populational interpretation*.

The startling nature of Boltzmann’s reinterpretation was due wholly to his great enthusiasm and deep understanding of the mechanistic explanation that Darwin had put forward regarding selection operating on the smallest of variations within a population of variants. Boltzmann was so entranced by Darwin’s ideas that he wrote in his 1905 book, *Populare Schriften*, “If someone asked me what name we should give this century, I would answer without hesitation that this is the century of Darwin.”

Ilya Prigogine (1984), Nobel Laureate in Chemistry, wrote that

Boltzmann was deeply attracted by the idea of evolution, and his ambition was to become the “Darwin” of matter. Boltzmann’s approach had astounding successes. It has left a deep imprint on the history of physics. The discovery of the quantum by Planck was an outcome of Boltzmann’s approach. I fully share the enthusiasm with which Schroedinger wrote in 1929 that ‘[Boltzmann’s] line of thought [Darwinism] may be called my first love in science. No other has ever thus enraptured me or will ever do so again.’

(Prigogine 1980).

Eighty years later, Shannon reinterpreted entropy for a third time, directly emulating Boltzmann, but this time, defining information as a metric of surprise rather than disorder. In that, a philosophical circle was completed. The simplest, most formal definition of the Darwinian paradigm is an infinitely iterated optimization algorithm. There is no sense of tautology in this definition. Rather, it is a simple thermodynamic process where the agent of selection (competition among the inevitable excess within a

bounded arena) operates to minimize the behavioral inappropriateness (surprise) of the trial variants, generation after generation, moving the evolving lineage across an adaptive topography to a point of optimization.

Shannon, too, defined an entropy, H' , but not as Boltzmann had done. Rather, his entropy was the weighted average of the unexpectednesses of the symbols within the symbol set. Although the story may be apocryphal, Shannon greatly worried about distorting Boltzmann's definition and addressed his concerns to John von Neumann. Von Neumann told him, "Don't worry. No one knows what entropy is, so in a debate you will always have the advantage."

Boltzmann's 1872 definition of entropy gave us for the first time a profound and measurable sense that time has an "arrow," in sharp contrast to the previous Newtonian physics, where reversibility is not only possible, but assumed. In doing so, Boltzmann simultaneously gave Darwinian evolutionary theory an even more fundamental philosophical footing. Boltzmann's entropy also provided the philosophical underpinning for the definition of the quantum, and thus all of quantum mechanics, and by further derivation, statistical mechanics.

Shannon's entropy very similarly equated the imposition of noise into the signalling channel as an increase in disorder, and thus allowed Shannon to define the additional qualities of *ambiguity* and *equivocation* in a transmitted message. These several properties became the philosophical bases for much of the error-correction, cryptography, and compression algorithmic theories that followed. Error-correcting codes add redundancy to a signal in order to insure proper reception, while in contrast, compression algorithms are intrinsically designed to remove as much of the redundancy from the transmitted signal as they possibly can, thereby maximizing the compressed signal's entropy.

The economist Kenneth Boulding once quipped that "Teaching is an [informational] entropy minimization process." In nature, selection is the

entropy minimizing process to an evolving lineage, and assumes the role of Boulding's teacher. The obvious result is that a good biologist can look at the face, gut, and feet of the animal he or she is studying, and in the absence of any other context, with some certainty specify the environment to which the animal has become adapted and its habits of life.

If we define *fitness*, not in the modern sense of *differential reproduction*, but as Darwin originally meant the term to be understood, as *appropriateness* to the tasks and circumstances at hand, then it is easy to see that evolutionary adaptation is equivalent to Boulding's entropy minimization. While this definition is impractical in any sense for field use, it is an enormously more accurate representation of the actual Darwinian physics in force.

Surprise is very often lethal in natural circumstances. The direct consequence of that lethality is that the qualities that evolve within the genomes of the competing phyletic lineages are increasingly accurate models of the frequencies and importances of the strings of symbols that the lineages are likely to encounter. Evolution operates wholly in a predictive mode, predicting the arrival of the next symbol, working as well as it can so as to nearly eliminate surprise.

If this line of thought is good for the goose, then it's just as good for the gaggle of species that comprise a community. Presence-absence matrices are the simplest, most basic form of ecological data we can obtain, but they are also among the most important. Merely knowing the presence or absence of a particular species at a site carries with it a great deal of information.

E. C. Pielou was almost certainly wrong when she wrote a quarter-century ago, "It cannot be too strongly emphasized that the fancied links between the information-theoretic concept of "information" and the diversity of an ecological community are merely fancies and nothing more." (Pielou 1976:9).

The informational processes that hold for evolution of lineages hold

for communities as well, and the apparent species diversities of communities are intimately tied to the heterogeneities of their respective landscapes, and thus to the landscapes' intrinsic information contents.

While we measure mass and energy flows in an ecosystem—in great part because we can—"life is information," and regardless of the practical difficulties that thought may hold for actual field work, it is a small, haunting chant that should constantly echo in the back of every field ecologist's mind.

Beyond the fundamental ideas of Boltzmann and Boole, Shannon's work built on the foundations of other researchers at Bell Labs such as Harry Nyquist, R. V. L. Hartley, Norbert Wiener, John von Neumann, and John Tukey. Indeed, it is surprising how much of this pioneering work in information theory, digital circuitry, computers, and telecommunications was done at only few locations: MIT, Bell Labs, and Los Alamos being the most prominent. The impetus for the extraordinary intellectual climate that developed these ideas was the urgency of World War II. Unfortunately, we are at the point in time when the last of these people are passing from this earth.

Nonetheless, the links between this highly practical work and living systems did not escape the participants. Shannon's doctoral dissertation was "*An algebra for theoretical genetics*," while Wiener became so entranced with the notion of negative feedback that in 1948 he published, *Cybernetics, or Communication and Control in the Animal and the Machine*. The same is as true of Alan Turing, John von Neumann, Leo Szilard, and Erwin Schroedinger.

What then of my original suggestion? If Shannon had never lived, would we know the same things we know now? Almost without doubt. There is an inevitability to the scientific process, just as there is to evolution. No better example of this exists than Shannon's information metric, surprise. Information exists in the novel stimulus, not the common one, and evolution discovered this very

basic truth half a billion years before Shannon. What neurophysiologists call *habituation* or *adaptation* occurs because of the evolved imposition of a negative feedback circuit on a neural pathway, deadening the pathway to the common stimulus, yet instantly waking it up when the unexpected occurs, giving it great weight.

Robert MacArthur's preface to *Geographical Ecology* is a bit of text that should be read and memorized by every young scientist, of any stripe: "To do science is to search for repeated patterns, not simply to accumulate facts. . . . Doing science is not such a barrier to feeling or such a dehumanizing influence as is often made out. It does not take the beauty from nature. The only rules of scientific method are honest observations and accurate logic. . . . Science should be general in its principles."

There may be a no more broadly repeated pattern than that of information, nor a more intriguing general philosophy than that there is a physics to information, and that we are beginning to understand it, simply because we first deduced how to manipulate it ourselves in the machines we build. Because of that, we find that we've come to understand the nature, mechanisms, and processes of life all the more deeply.

Literature cited

Boltzmann L. 1872. Weitere studien uber das warmegleichgewicht unter gasmolekullen. Wien Berichten **66**:275–370.

Boltzmann L. 1905. Populare schriften. Leipzig. [English translation: Reidel, Dordrecht, Germany/ Boston, Massachusetts, USA. 1974]

Clausius R. 1865. Ueber verschiedene fur die anwendung begueme formen der hauptgleichungen

der mechanischen warmetheorie. Annalen der Physik **125**:353–400.

Dawkins, R. 1986. The blind watchmaker: why the evidence of evolution reveals a universe without design. W. W. Norton, New York, New York, USA.

Eigen, M., with R. Winkler-Oswatitsch. 1992. Steps toward life: a perspective on evolution. Oxford University Press, Oxford, UK.

Johnson, G. 2001. Claude Shannon, mathematician, dies at 84. New York Times, 27 February 2001.

Lewontin, R. C. 1974. The genetic basis of evolutionary change. Columbia University Press, New York, New York, USA.

MacArthur, R. H. Geographical ecology: patterns in the distribution of species. Harper & Row, New York, New York, USA.

Pielou, E. C. 1975. Ecological diversity. John Wiley & Sons, New York, New York, USA.

Prigogine, I. 1980. From being to becoming: time and complexity in the physical sciences. W. H. Freeman, San Francisco, California, USA.

Prigogine, I., and I. Stengers. 1984. Order out of chaos. Bantam Books, New York, New York, USA.

Shannon, C. E. 1948. A mathematical theory of communication. The Bell System Technical Journal **27**:379–423 (July), **27**:623–656 (October).

Wiener, N. 1948. Cybernetics, or communication and control in the animal and the machine. MIT Press, Cambridge, Massachusetts, USA.

Wirt Atmar
AICS Research, Inc.
Las Cruces, NM 88003-4691
E-mail: atmar@aics-research.com,
atmar@fieldmuseum.org