Section One — Information Theory & Biology: Introductory Comments

Robert J. Marks II — Section Chairman

All agree there is information in biological structure and function. Although the term *information* is commonly used in science, its precise definition and nature can be illusive, as illustrated by the following questions:

- When a paper document is shredded, is information being destroyed? Does it matter whether the shredded document is a copy of an un-shredded document and can be replaced?
- Likewise, when a digital picture is taken, is digital information being created or merely captured?
- The information on a DVD can be measured in bits. Does the amount of information differ if the DVD contains the movie *Braveheart* or a collection of randomly generated digital noise?
- When a human dies, is experiential information lost? If so, can birth and experience create information?
- If you are shown a document written in Japanese, does the document contain information whether or not you know Japanese? What if instead, the document is written in an alien language unknowable to man?

The answers to these questions vary in accordance to the information model used. However, there are properties of information common to all models. As noted by Norbert Weiner [1, 2], the father of cybernetics:

> "Information is information, neither matter nor energy."

Information can be written on energy. Examples include wireless electromagnetic waves and audio waves that carry the content of conversations. As is the case with books and DVD's, information can also be etched onto matter. But energy and matter serve only as transcription media for information. Information can also reside in structure or phenomena. Varying degrees of information are available in nature. A bacterium obviously contains more information than a grain of sand. Information can be extracted from inspection of information-rich sources. The idea for Velcro came from close examination of burrs stuck to the clothes of a Swiss engineer after a hunting trip [3]. The function of the human eyelid was the inspiration for invention of the intermittent windshield wiper [4]. *The IEEE*

Computational Intelligence Society [5], a professional electrical and computer engineering organization,[1] has as its motto, "Nature inspired problem solving." The implication is that structure in nature, when examined, can be a rich source of information applied to engineering. Unlike mass and energy in physics, a single model or definition of information does not exist. Claude Shannon recognized his theory was not the last word in the mathematical modeling of information [6]:

> "It seems to me that we all define 'information' as we choose; and, depending upon what field we are working in, we will choose different definitions. My own model of information theory... was framed precisely to work with the problem of communication."

Shannon Information

Because of its widespread application and depth of mathematical rigor, the most celebrated information model is Shannon information theory. In an astonishing 1948 paper [7], Claude Shannon single-handedly founded a discipline still celebrated today by professional organizations such as the *IEEE Information Theory Society* who has published The IEEE Transactions on Information Theory since the mid-1950's. Shannon's original paper is remarkable. The word *bit*, a contraction of *binary digit*, was first used in this paper.[2] To show that continuous time signals could be represented by discrete time samples, Shannon discussed the sampling theorem[3] that is today a universal staple of undergraduate electrical engineering curricula [9], and dictates how many discrete samples must be captured on DVD's and digital images to faithfully reconstruct continuous time audio signals and images [8, 9]. A relationship between average information and thermodynamic entropy was established by Shannon. In one of the most important applied mathematical results of the twentieth century, Shannon also showed that errorless communication was possible over a noisy channel. Forty five years later, *turbo codes* for the first time came very close to achieving the errorless communication bounds predicted by Shannon [10].

A fundamental contribution of Shannon's paper is a mathematical definition of information. Two axioms are foundational to Shannon information.

---

[1]IEEE, the *Institute of Electrical and Electronic Engineers*, is the world's largest professional society. In 2010, there were 382,400 members.

[2]Shannon credited John W. Tukey, a fellow Bell Labs researcher, with coining the word.

[3]I wrote an entire book dedicated to this topic [8], only one of the amazing contributions of Shannon's paper.

1. As the probability of an event increases, the amount of information decreases. There is little or no information in the statement that the sun will rise tomorrow morning. The probability of the event is nearly one. On the other hand, the event of the sun going supernova tomorrow has a miniscule almost zero probability. Being told the sun is going supernova tomorrow conveys much information.
2. Information of two disjoint events should be additive. That is, if the word "stuttering" conveys information I_1 and "professor" conveys information I_2, then "stuttering professor" should convey information $I_1 + I_2$.

If p denotes the probability of an event, the definition that satisfies both of these axioms is[4]

$$I = -\log_2 p.$$

As required by the first axiom, information increases as probability decreases. If two disjoint (statistically independent) events have probabilities p_1 and p_2, then the probability of both events is $p_1 p_2$ with information $I = -\log_2 p_1 p_2 = I_1 + I_2$ where $I_1 = -\log_2 p_1$ and $I_2 = -\log_2 p_2$. The additivity axiom is thus satisfied.

The base of the log in the definition of Shannon information is arbitrary and determines the units of information. If base 2 is used, then the unit of information is a *bit*. If a fair coin is flipped 6 times, we can say there are six bits of information generated since the probability of generating a specific sequence, say HTTHH, is

$$p = \left(\frac{1}{2}\right)^6 = 2^{-I}.$$

The bit can be viewed as probability measured in coin flips. Ten bits, for example, corresponds to successfully forecasting the results of ten coin flips. Pioneering application of Shannon information theory to biology includes the work of Thaxton, Bradley & Olsen [12] and Yockey [13, 14]. There are limitations to Shannon information. Isolated from context, Shannon information measure is divorced from meaning. A *Braveheart* DVD can contain as many bits as a DVD filled with random noise. When applying Shannon information, care must be taken to recognize this property and, if meaning is applicable, to make clear the connection.

[4]Use of the log to measure information dates to 1928 when Ralph Hartley noted that "...our practical measure of information [is] the logarithm of the number of possible symbol sequences." [11] This is equivalent to Shannon information when all symbol sequences are equally probable.

Solomonov-Kolmogorov-Chaitin Information

Shannon information is motivated by communication. *Algorithmic information theory*, also called Solomonov-Kolmogorov-Chaitin information after the three men who independently founded the field[5] [15–22], is a topic in the field of computer science. Whereas Shannon information deals with probability of future or unknown events, algorithmic information deals largely with the complexity of existing structure. To what degree can a thick book, say the KJV Bible, be compressed? The length of the shortest computer program to generate KJV Bible is dubbed the Chaitin-Kolmogorov complexity of the book.[6] A repeated sequence 010101010... for a billion bits has low complexity. The computer program is "Repeat 01 a half billion times." A billion bits generated by repeated flips of a fair coin, on the other hand, is almost certainly incompressible. The shortest program to print the sequence must then contain the sequence, "Print '0110100010....'."

An implication of the word *information*, when used conversationally, is the communication of meaning. Algorithmic information theory's measure of complexity suffers from the same problem as Shannon's model—it does not inherently capture the meaning in the information measured [23]. A digital image of a Caribbean sunset can have the same Chaitin-Kolmogorov complexity as an unfocused image of correlated noise.

The Meaning of Information

Meaning in information is captured by the concept of *specified complexity* popularized by Dembski [24, 25]. The idea can be illustrated using the English alphabet [12]. The phrase

OVER AND OVER AND OVER AND OVER AND OVER AND
OVER AND OVER AND OVER AND OVER AND OVER AND

has specific meaning but has a low Chaitin-Kolmogorov complexity. A program can read "Repeat 'OVER AND' ten times." The phrase

HSUEX SHDF OSJ HDFN SJABXMJ SHBU SZJLK QPRQZ HASKS
FPSCSJSJAA PJKAO DFAJ AJDFHFQWSALA DAFL V AZQEF

[5]Chaitin, born in 1947, was still a teenager when his first groundbreaking work was published in 1966.

[6]The minimum program depends on the computer program used, but the measure from computer to computer varies only by an additive constant.

is complex. The program for this phrase would be "Print HSUEX SHD... ZQEF". This is about the same size of the phrase itself. The phrase however, has no specified meaning. Next, consider the Bob Dylan lyrics[7]

I ASKED FOR SOMETHING TO EAT IM HUNGRY AS HOG SO I
GET BROWN RICE SEAWEED AND A DIRTY HOT DOG.

This sequence of letters, display both a specified meaning and high complexity.

Leslie Orgle notes, regarding the requirement of specified complexity in life:

> "Living organisms are distinguished by their specified complexity. Crystals such as granite fail to qualify as living because they lack complexity; mixtures of random polymers fail to qualify because they lack specificity." [26]

Orgle's statement was independently observed by Yockey and Wickens [12]. Other models of information include universal information [1], functional information [23, 27, 28], pragmatic information [29] and evolutionary informatics [30–32]. Except for functional information, all of these models are addressed in this section.

Papers

The papers in this section on *Information and Biology* fall into three distinct categories.

1. *Information Theory Models*

How can information be modeled to reflect the information residing in biological systems? **Gitt, Compton and Fernandez** [1] define *universal information* as; "A symbolically encoded, abstractly represented message conveying the expected action and the intended purpose." They then show how universal information is resident in biological systems. **Dembski *et al.*** [43] build on the theory of evolutionary informatics [30–32] by developing a generalized search methodology. Using conservation of information ideas popularized by the No Free Lunch theorem [25], evolutionary search is shown to produce no *active information*. The difficulty of the search at hand, measured by *endogenous information*, can be simplified only by access to some source of information. **Oller**'s *pragmatic*

[7] "On the Road Again" by Bob Dylan.

information [29] refers to the content of valid signs — the key that unlocks language acquisition by babies and ultimately leads to human communication through language. Oller shows this same measure is required for "codes" in genetics, embryology, and immunology to work.

2. *Limitations of Evolutionary Models*

A colleague of mine visiting my office noticed my computer buzzing away. When he asked what I was doing, I replied "running a self-organizing evolutionary program." In mocked astonishment, he queried "That's exciting! When will it be able to talk?" The truth in this quip is that evolutionary systems often hit a point after which no further improvement is observed. Behe [37], who coined the phrase *edge of evolution*, documents that biological evolution can also develop to a point where no other improvement is observed. In such case, specified complex information is bounded. **Basener** [38] proves such a ceiling of performance exists in many evolutionary processes. Specifically he finds; "In an evolutionary system driven by increasing fitness, the system will reach a point after which there is no observable increase in fitness." Schneider's *ev* [39] and Avida [40] computer programs that purport to demonstrate biological evolution obey the criteria necessary for Basener's result to apply. No matter how long they run, neither program will ever learn to talk. **Ewert *et al.*** [41] demonstrate that TIERRA, Thomas Ray's attempt to simulate a Cambrian explosion on the computer, also hits Basener's ceiling. Although TIERRA demonstrates fascinating and unexpected behavior, interesting innovations consistently arise only from loss of function. This same phenomenon in biology is reported by Behe [37]. **Montañez *et al.*** [42] assess the probability of information being increased via random mutations within a genome. They show that the probability of improvement drastically diminishes as the number of overlapping codes increases and to the extent that the DNA sequence is already near its optimum.

3. *Thermodynamics, Entropy and Information*

Both information theory and thermodynamics share the concept of entropy referring to maximum disorder and uncertainty. Recognizing that life does not conform to thermodynamics' demand for ever increasing disorder, Erwin Schrödinger coined the term *negentropy* (negative entropy) to apply to life. What is the source of negentropy?

Sewell [35] shows that the decrease of entropy within a non-isolated system is limited not by "compensating" entropy increasing outside the system, but by the type and amount of entropy exported through the boundary. Thus, in open systems, information increases are limited by the information entering through the boundary. In other words, it is not true that *anything* can happen in an open system [36]. **McIntosh** [33] carefully argues that the laws of thermodynamics do not permit the rise of functional devices ('machines') just by the flow of energy into a non-isolated system. Free energy devices available to do useful work are a product of intelligence. If one then considers information itself, one then finds that rather than matter and energy defining the information sitting on the polymers of life (a view held by many today), McIntosh posits that the ***reverse*** is in fact the case. Information has its definition outside the matter and energy on which it sits, and furthermore constrains matter/energy to operate in a highly non-equilibrium thermodynamic environment. He then outlines principles of information interaction with energy and matter in biological systems [34].

A Final Thought

Much work remains on development of a concise mathematical model of information applicable to biological systems. Some physicists have argued that all of the information required for the observable universe, including physical laws and the prescription for life, was created through the Big Bang. The authors of this section appear to unanimously disagree with such an assertion.

References

1. Gitt W, Compton R, Fernandez J (2013) Biological Information — What is It? In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 11–25.
2. Wiener N (1968) Cybernetics: or the control and communication in the animal and the machine. Technology Press, MIT.
3. Freeman A, Golden B (1997) Why Didn't I Think of That: Bizarre Origins of Ingenious Inventions We Couldn't Live Without. John Wiley & Sons, New York. ISBN 0471165115
4. Seabrook J (1993) The Flash of Genius. The New Yorker, 11 January 1993.
5. IEEE Computational Intelligence Society, http://www.ieee-cis.org/

6. Mirowski P (2002) Machine dreams: economics becomes a cyborg science. Cambridge University Press, Cambridge.
7. Shannon CE (1948) A Mathematical Theory of Communication. Bell System Technical Journal 27: 379–423, 623–656.
8. Marks II RJ (1991) Introduction to Shannon Sampling and Interpolation Theory. Springer-Verlag, New York.
9. Marks II RJ (2008) Handbook of Fourier Analysis and Its Applications. Oxford University Press, Oxford.
10. Guizzo E (2004) Closing in on the perfect code. IEEE Spectrum, March 2004.
11. Hartley RVL (1928) Transmission of information. Bell System Technical Journal 7 (3):535–563.
12. Thaxton CB, Bradley WL, Olsen RL (1992) The Mystery of Life's Origin: Reassessing Current Theories, 2nd print edn. Lewis and Stanley. [http://TheMysteryOfLifesOrigin.org/]
13. Yockey HP (1992) Information Theory and Molecular Biology. Cambridge University Press, Cambridge.
14. Yockey HP (2005) Information Theory, Evolution, and The Origin of Life. Cambridge University Press, Cambridge.
15. Chaitin GJ (1966) On the length of programs for computing finite Binary sequences. J. Assoc for Computing Machinery 13 (4):547–569.
16. Chaitin GJ (1966) On the length of programs for computing finite binary sequences by bounded-transfer Turing machines. AMS Notices 13: 133.
17. Chaitin GJ (1966) On the length of programs for computing finite binary sequences by bounded-transfer Turing machines II. AMS Notices 13: 228–229.
18. Chaitin GJ (1966) On the length of programs for computing finite binary sequences. Journal of the ACM 13: 547–569.
19. Kolmogorov AN (1965). Three approaches to the definition of the quantity of information. Problems of Information Transmission 1: 3–11.
20. Kolmogorov AN (1968) Logical basis for information theory and probability theory. IEEE Trans. Inform. Theory, vol. IT-14: 662–664.
21. Solomonoff RJ (1960) A preliminary report on a general theory of inductive inference, Tech. Rept. ZTB-138, Zator Company, Cambridge, MA.
22. Solomonoff RJ (1964) A Formal Theory of Inductive Inference. Information and Control 7 (1):1–22 & 7(2): 224–254.
23. Szostak JW (2003) Functional information: Molecular messages. Nature 423(6941): 689.
24. Dembski WA (2002) Can evolutionary algorithms generate specified complexity? In: Gregersen NH (ed) From Complexity to Life: On The Emergence of Life and Meaning. Oxford University Press, USA.
25. Dembski WA (2007) No Free Lunch: Why Specified Complexity Cannot Be Purchased without Intelligence. Rowman & Littlefield, Lanham, MD.
26. Orgel LE (1973) The Origins of Life. John Wiley and Sons, Inc., New York.

27. Durston KK, Chiu DKY, Abel DL, Trevors JT (2007) Measuring the functional sequence complexity of proteins. Theoretical Biology and Medical Modelling 4:47 doi:10.1186/1742–4682–4-47 [http://www.tbiomed.com/content/4/1/47]
28. McIntosh AC (2006) Functional Information and Entropy in living systems. In: Brebbia CA, Suchrov LJ, Pascolo P (eds) Design and Nature III: Comparing Design in Nature with Science and Engineering. WIT Press, UK.
29. Oller JW (2013) Pragmatic Information. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 64–86.
30. Dembski WA, Marks II RJ (2009) Conservation of Information in Search: Measuring the Cost of Success. IEEE Transactions on Systems, Man and Cybernetics A, Systems and Humans 39 (5):1051–1061.
31. Dembski WA, Marks II RJ (2010) The Search for a Search: Measuring the Information Cost of Higher Level Search. Journal of Advanced Computational Intelligence and Intelligent Informatics 14(5):475–486.
32. Dembski WA, Marks II RJ (2011) Life's Conservation Law: Why Darwinian Evolution Cannot Create Biological Information. In: Gordon B, Dembski W (eds) The Nature of Nature (Wilmington, Del.: ISI Books, 2011) pp.360–399.
33. McIntosh AC (2013) Information and Thermodynamics in living systems. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 179–202.
34 McIntosh AC (2009) Information and Entropy — Top-down or bottom-up development in living systems? Int. J. of Design & Nature and Ecodynamics 4(4):351–385.
35. Sewell G (2013) Entropy, Evolution and Open Systems. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 168–178.
36. Sewell G (2001) Can ANYTHING Happen in an Open System? The Mathematical Intelligencer 23(4): 8–10.
37. Behe MJ (2007) The Edge of Evolution: The Search for the Limits of Darwinism. Free Press, New York.
38. Basener WF (2013) Limits of Chaos and Progress in Evolutionary Dynamics. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 87–104.
39. Montañez G, Ewert W, Dembski WA, Marks II RJ (2010) Vivisection of the ev Computer Organism: Identifying Sources of Active Information. BIO-Complexity 2010(3):1–6.
40. Ewert W, Dembski WA, Marks II RJ (2009) Evolutionary Synthesis of Nand Logic: Dissecting a Digital Organism. Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics. San Antonio, TX, USA — October 2009, pp. 3047–3053.

41. Ewert W, Dembski WA, Marks II RJ (2013) Tierra: the character of adaptation. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 105–138.
42. Montañez G, Marks II RJ, Fernandez J, Sanford JC (2013) Multiple overlapping genetic codes profoundly reduce the probability of beneficial mutation. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 139–167.
43. Dembski WA, Ewert W, Marks II RJ (2013) A General Theory of Information Cost Incurred by Successful Search. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 26–63.