The Unfolding of "Information" in Big History

Ken Solis

Abstract

This essay's central thesis is that information and its "flows" are just as crucial as energy flow densities for the realization of increasingly complex systems over the course of big history. In fact, it is the requisite interplay between at least these two phenomena that make complexity possible. As with any endeavor of a philosophical or scientific nature, definitions are a crucial beginning point for building any argument. Hence, critical definitions will form an important basis for the content of this essay. If we assert that information plays just as essential a role as energy flows do for the realization of complex systems, then we must also propose the role information plays on a more basic physical level to support this contention. After all, complex systems don't "spring fully formed from the head of Zeus like Athena." Finally, and perhaps most importantly, we must have a better intuitive grasp of what information is or at the very least, what it does. As an example, physicists still don't know what "energy" fundamentally is, just what it does, e.g., "energy is the capacity to do work."

Keywords: Information theory.

Introduction – What *is* "Information?"*

The challenge posed in defining "information" is perhaps best reflected in a quotation from Claude Shannon (1916-2001), the widely acknowledged "father" of information theory:

> The word "information" has been given different meanings by various writers in the general field of information theory. It is likely that at least a number of these will prove sufficiently useful in certain applications to deserve further study and permanent recognition. *It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field.* (Shannon, 1993, p180)

Inadvertently supporting his contention, luminaries from different fields of study have offered varied definitions of information. The following are just a few examples:

> "The ability to distinguish reliably among possible alternatives." Claude Shannon, founder of information theory. (Schumacher, 2015, p6)

> "It from bit." John Wheeler, renowned 20th century physicist. (Wheeler, 1990)

> "A difference that makes a difference." Gregory Bateson, anthropologist (Bateson, 1972)

> "The difference between maximum entropy and actual entropy." David Layzer, astrophysicist (Layzer, 1990, p28)

> "If information *is fundamentally relational*, then it makes sense that it is limited by surface area." Benjamin Schumacher, quantum information pioneer and last graduate student of John Wheeler. (Schumacher, 1915b, p550)

I would suggest that Wheeler's and Bateson's definitions are pithy observations of what information *does*. In Wheeler's "It from bit" claim, he is stating that information results in the genesis of the "things" or structures of the universe as particles and even the entire universe developed from patterned relationships. Bateson similarly is making the claim that information essentially causes or results in

---

* My thanks to Bob Doyle, PhD, a former colleague of Eric Chaisson, and fellow protégé of astrophysicist, David Layzer. While Dr. Chaisson analyzed much on the energy flow densities of complexity, Dr. Doyle was more interested in information and philosophy (see www.informationphilospher.com). I am deeply in his debt for being among the first to make me better appreciate that my intuition about information's importance had true merit and deepened my appreciation for this topic.

processes, or both. Layzer's definition, proposes a measure for the amount of information of a system, including that of the universe (Layzer, 1990, p28). Schumacher's definition, however, gets to the core of what information *is*. Similarly, Terrence Deacon, a neuro-anthropologist at the University of California - Berkeley, came to the same conclusion as Schumacher and matter-of-factly states, "information is about *the relationship of something to something else*." (Deacon, 2010, p159). Independently, but admittedly at a later date than both Schumacher and Deacon offered their definitions, I proposed that information is: "The relationships of entities in space-time." I added the phrase, "in space-time," because the scientific community has not come to a consensus as to what happens to information that enters black holes, where the known laws of normal space-time might not apply (Seife, 2007, pp 230-40).

Some Examples that Illustrate Why Information = Relationships

Here are at least a few arguments as to why information is fundamentally relational in nature:

- Imagine a very simple piece of information: for example, what is the location of a singular particle in otherwise empty space? Ultimately, in a 4 dimensional universe, you would give its coordinates (e.g., x_1, y_1, z_1, t_1) in relationship to something else, like the boundaries of space, its center, or perhaps another entity, plus some relation for "time," e.g., years since the Big Bang, the, the founding of Rome, the birth of Christ. In short, we also need a x_0, y_0, z_0, t_0 as reference (or relational if you prefer) coordinates. If a particle is simply present in a boundless, infinite space-time without any relationship to something else, you will not be able to give any information about its coordinates or even determine if it is stationary or moving.
- As Benjamin Schumacher points out in his Great Courses lecture series, *The Science of Information*, cosmic information has been postulated to reside on the surface area of the incredibly miniscule piece of space, about 7 x 10^{-70} m^2. Similarly, information is believed to exist on the surface or event horizon of a black hole (Schumacher, 2015a, p284). He pointed out that this makes sense, because it is the surfaces of the smallest possible "units" of space, or black holes that interacts or has relationships with the rest of the universe.
- Scientists are having a difficult time engineering a robust quantum computer because the "qubit" particles have to remain in a superpositional quantum state for it to work correctly. As soon as a qubit interacts (has a relationship with another part of the universe), it "decoheres" into a classical bit of information, i.e. it becomes a "1" or "0" rather than something, somewhere in between (Schumacher, 2015b, p499).
- .At the very beginning and end of big history - after the universe's possible future heat death - very little or no information will exist as structures even as simple as hadrons (e.g., protons, neutrons) might break down into random radiation (no patterns or structures), and no energy differentials are present to drive processes (Christian, 2011, p489). We will discuss information's relationship to entropy in greater detail later.

Despite the foregoing arguments and Deacon's implication that information's equivalency to relationships is obvious, not everyone has come to the same conclusion. After all, Claude Shannon, the "father" of information theory himself did not believe that any single concept of information would be satisfactory. However, despite his genius, Shannon may have been partially mistaken in this regards, if only because as Deacon again points out, *"This term is used to talk about a number of different kinds of relationships, and often interchangeably without discerning between them"* (Deacon, 2011,

p152). I concur and believe that we must start with understanding and parsing at least the primary types of information before we can proceed forward.

Syntactical Information is the fundamental type of information that underlies all others and is, therefore, pervasive. Despite this basis, it is less intuitive to many people because our colloquial use of the term more commonly refers to other types of information which I will discuss shortly. Syntactical information, as I defined earlier, is about relationships. Hence, it can also be conceived as any pattern (a more static relationship) or process (a more dynamic relationship) of or between "things" that compose the constituents of the universe. As a parallel, in linguistics "syntax" refers to the rules of how different kinds of words (nouns, verbs, adjectives, etc.) are generally ordered for a particular language. Claude Shannon, a brilliant mathematician and engineer, was also amongst the first to understand that the task of designing communication technology was a challenge of accurately transmitting only syntactical information from one place or time to another. The *meaning* of the message is not relevant for engineering communication technology (Shannon, 1948). Engineers just need to develop the tools and processes to faithfully transmit patterns across time and space. The patterns can be variations in an electrical discharge as with Morse codes, variations in radio frequency amplitude as with some radio communications, alternating bands of black and white as with bar codes, and so on. Furthermore, information can transmute from one type of pattern to another and even between different kinds of medium.   An oversimplified example can illustrate this chain of informational transmutations: A TV camera detects the photon emission pattern of wherever the lens is pointed; this photon pattern is converted to a pattern of electrical discharges, which is later converted to radio waves that are transmitted to a satellite; the satellite converts those waves to electrical patterns again, and then back to radio waves to be transmitted back to an antenna back on Earth. . . This general process of informational transmutations continues until your own TV transmits photons in a pattern very similar to the original source to your eyes. Still, this process continues via your central nervous system until "you" apprehend a satisfactory reproduction of the original image that might have occurred thousands of miles away!

Fundamentally, syntactical information also existed long before language and even life began – essentially since the Big Bang. *Natural* syntactical information is driven by many forces and processes, most importantly the three fundamental forces of the universe: (1) electrical weak force - often separated into electromagnetism and the nuclear weak forces, (2) strong nuclear force, and (3) gravity. Hence, the strong nuclear force causes quarks to relate to each other to form protons, neutrons, other particles, and atomic nuclei. The electrical weak force causes electrons to relate to nuclei to form atoms and molecules. Gravity, meanwhile, causes atoms, nuclei, and other subatomic particles to come together to form massive structures like planets, stars, and galaxies. Examples of *processes* driven by these same forces include nuclear fusion, photon emission, and the orbits of planets around a star, respectively. Other cosmic ingredients like dark matter and possibly dark energy are also relevant in the working of the cosmos although their nature is not understood at this time. The only thing we know about them is informational -- that they cause "ordinary" matter and energy to relate differently than can be accounted for by the known constituents and forces of nature. For example, dark matter was "discovered" when astronomers determined that there is not sufficient ordinary matter to cause galaxies and galaxy clusters to stay together. Dark energy was "discovered" when astronomers determined that galaxies are moving away from each other faster than can otherwise be explained. It is also important to realize that we do not even know if dark matter is matter, or that dark energy is energy -- the names are simply placeholders.

Semantical Information is information which has been processed by an *agent* so that it now has purpose or meaning for that or another agent. Per the Merriam-Webster dictionary, an *agent* is, "one that acts or exerts power." The question of when and how an "agent" becomes extant would easily consume an entire area of speculative science and philosophy, as can questions surrounding semantics itself (Floridi, 2011). For the purposes of this essay, however, an agent is a living organism that is able to act. (One might argue that artificial intelligence can also process information so that it can become semantical, but I will leave that philosophical discussion to the side.) A typical example of semantic information is the statement: "Please arrive at the Melrose Diner at 5 p.m. today." The letters are ordered in such a way as to form words that have meaning(s). The words are also chosen and arranged per the general rules of a language (syntax), so that one agent purposefully informs another that they should meet at a certain place and time. The import of the sentence above typifies what most people imagine when we say the word "information," although we would also include data transmitted and received by the internet, smart phones, television, books, and so on. Per the definition of semantic information, however, even sunlight shining on a "simple" plant has semantic informational content. The sunlight has meaning or purpose for the plant and might cause it to turn its leaves toward the sunlight to gather more energy for sustenance – even if the plant is not consciously aware of its actions.

Also, agents can create *artificial* syntactical information instinctually, by intent, or accidentally, e.g., the utterance of sounds, a bear leaving claw marks on a tree, the release of pheromones, and so on. That artificial information can then become semantical for itself or to another agent. After all, the strips of missing bark on a tree are just that on the purely physical descriptive level. To a wandering bear, however, the missing strips of bark informs it that it has entered the territory of another bear.

Finally for our discussion, "novel," a.k.a. "pragmatic," information is that which provides new data to an agent and, thus, makes them aware, or at least more certain of a relationship of which they were not previously aware or about which they were uncertain. A classic example of novel information is when two lanterns were hung in Boston's Old North Church tower on April 18, 1775 c.e. to inform Paul Revere and others that the British were traveling by sea rather than by land to reach Lexington. The hanging of a chosen number of lanterns in the tower exemplifies what Shannon describes with his definition of information as "The ability to distinguish reliably among possible alternatives." In this case, Paul Revere had his uncertainty reliably reduced as to which route the British troops would travel. Novel information is a subset of semantic information because it also requires an agent, and not all semantic information provides new information to it, e.g., "the sun came up in the east this morning."

Information and Microstates

It might seem strange to consider an atom or a planet, for example, to in any way consist of information, but more sensible for them to be *described* by information, e.g., the planet's circumference is X kilometers, its mass is Y kilotons. However, physicists consider parameters like size, mass, temperature, and so on to be macrostate (~overall) properties of a system. According to thermodynamics, one of the principle disciplines of physics, a macrostate of a system is in turn determined by a corresponding number of possible microstates, which is how a system's microscopic constituents are arranged, interact, and behave. Temperature, for example, depends on the average kinetic energy of a system's molecules. A system with fast moving or quickly vibrating molecules has a higher temperature than one with slower moving or vibrating molecules. A system with closely packed molecules will have a higher density than one with more loosely packed molecules of the same kind.

Hence, how a system's microscopic constituents are arranged and interact with one another are its "microstate," and is synonymous with the relationships of its constituents, i.e. the information exchanged by its microscopic constituents in turn causes a system's macroscopic properties. It is this very type of information that is invoked in the famous Thorne-Hawking-Preskill wager: "Is the information that crosses the event horizon of a black hole (e.g., the mass, composition, particle motions, etc., of a star) lost to the universe forever or is it somehow preserved?" Although Stephen Hawking has contended that information is preserved, not everyone is convinced, including professors Thorne and Preskill (Gleick, 2011, pp358-9). Perhaps an easier to comprehend and accept example of a natural syntactical informational structure is the DNA molecule of a chromosome. Not only is DNA arranged in very particular patterns, it "contains" much of the information needed for the incredibly complicated functioning and reproduction of entire living organisms. (A little more on that later.)

Information – the Obverse of Entropy

As alluded to earlier, the flip side of a high degree of order or informational content is, roughly, a high degree of *dis*order. In thermodynamics, the degree of disorder or a system is often referred to as entropy. More technically, entropy is, "the log of the number of a system's microstates (or possible microscopic combinations) that can represent a macrostate (its large scale properties) (Stone 2015, p173). The fundamental formula for measuring entropy as described by the Austrian physicist, Ludwig Boltzmann (1804-1906) is: "$S = k \log W$." "S" is entropy, "k" is a constant, and "W" is the possible number of microstates that are possible for a particular macrostate of a system. (If needed, see the side bar for a brief review of log functions.)

Very similarly, the simplest expression of the amount of information in a message is $H = -k \log M$, where "H" is the amount of information, "$-k$" is a constant, and "M" is the probability of a message. Note that the equations given above are the simplest expressions of a measure of entropy and information, respectively. Slightly more extended formulas that cover more situations are typically used in the respective sciences, but the parallels between these formulas remain consistent, nevertheless. Also note that the values of "k" or "$-k$" do not mitigate the parallel either.

The similarity of the equations for information and entropy is not coincidental as was noted very early by Claude Shannon and other scientists. In fact, information theory was eventually used to successfully solve a century's old riddle in thermodynamics regarding a possible loophole to the second law of thermodynamics, called "Maxwell's Demon." In brief, in 1867 the famous physicist, James Clerk Maxwell (1831-1879), proposed a hypothetical way for a microscopic super-being to violate the second law of thermodynamics - which states that entropy of an <u>isolated</u> system always remains the same or increases. Attempts to disprove Maxwell's Demon using arguments from various areas of physics failed. By 1961, Rolf Landauer (1927-1999) proposed how information theory shows that the Demon cannot thwart entropy, and Charles Bennett (b. 1943) proved this conclusively in 1982. In the end they showed that it was the inevitable *erasure* of information that *must* incur energy costs, and hence would increase the entropy of any process in an isolated system (Seife, 2007, pp80-7). The main point is that information has been demonstrably proven to be essentially the flip-side of entropy or another aspect of entropy as some prefer to view it.

The laws of thermodynamics are also considered by many to be the most inviable laws in all of physics. The second law of thermodynamics is considered especially unassailable by physicists, including astrophysicist and philosopher Sir Arthur Eddington (1882-1944) who said, "The law that entropy always increases – the second law of thermodynamics – holds, I think the supreme

position among the laws of Nature." (Seife, 2007, p34). Before proceeding, it should be restated that entropy of any *isolated* system must remain the same or increase – it can "never" decrease (a little more on "never" a bit later). Hence, it is possible to decrease the entropy of one part of a system, as long as that decrease is more than offset by an increase in the system's overall entropy. A star, for example, appears to decrease entropy or disorder when gravity causes the particles of a nebula to form a compact more organized sphere. That decrease in entropy is more than offset, however, by the subsequent emission of photons, neutrinos, and other particles back out into space (Chaisson, 2001, p73).

The apparent "force" of entropy actually stems from raw statistical power. To illustrate, let us look at a functional car as an example of a small system with low entropy. As already noted, entropy is the log of the number of microstates (e.g., assemblages of car parts) that can represent a system's macrostate (overall properties of a *functional* car). According to Toyota, a car is comprised of about thirty thousand (3×10^4) parts (see http://www.toyota.co.jp/en/kids/faq/d/01/04/). For a car to work properly, the parts must relate to each other in a limited number of ways. You could, for example, change the seats around, or switch the lug nuts and still have a functioning car. Let's be charitable and say that there are about 10^5 ways to assemble a car's parts so that it is still in a functional macrostate. While one hundred thousand ways to assemble a working car from thirty thousand parts might seem like a large number of permutations, the *possible* number of ways to arrange over thirty thousand parts is an incredibly vast number. The mathematical formula for the number of possible permutations is a factorial of thirty thousand, or 30,000 x 29,999 x 29,998 x 29,997 x . . . 2 x1. Consider this: if there were only sixty parts to a car, the number of possible permutations for arranging the parts is 8.32×10^{81}, about the same as the number of particles in the observable universe (Seife, 2007, p65)! No wonder it is far easier to

take a car apart than it is to put it back together. Also, mathematically, a functioning car has very low entropy ($S = k \log 10^5$) when compared to a disassembled car with scattered parts ($S = k \log 10^{>>81}$).

It's important for us to note that entropy technically does not absolutely preclude the incredibly remote possibility of a car spontaneously forming all the right relationships to form a working car again. If the parts were floating in space in a box to keep the parts in close proximity, and an energy source was available to tighten bolts, etc., it is hypothetically possible for the car to come together again spontaneously to make a functioning car because the underlying physics are reversible. However, the statistical chance of this occurring is so miniscule, that the universe would long expire before there is a reasonable chance for this phenomenon to occur. Hence, the law that entropy "always" increases for a process, or a car "never" reassembles itself has a chance of being wrong, but it is statistically so miniscule that for all intensive purpose we can still say "always" and "never."

To further illustrate the mathematical statistical power of entropy, note that I only counted the large scale parts of the car and not the incredibly vast number of atoms and molecules that make up the car, and are also prone to other forms of disassociation from oxidation, ultraviolet light degradation, thermal motion, quantum fluctuations, etc. For example, if you included just the number of molecules in 3.5 tsp of water, the number of possible permutations for those molecules of water is over 10 to the 10th power with 24 zeros after it. Now, imagine the number of molecules that constitute a car versus a teaspoon of water, and the number of possible permutations for the car molecules and atoms become incalculably enormous – the vast, vast majority being in a "nonfunctional-car macrostate." Nevertheless, somehow the forces inherent in our universe, made ever increasingly informationally rich structures and processes extant. Mathematically, $H = -k \log$

M, where M, is the number of possible yes or no messages that define a structure or process (H). The more a structure's or process's constituents must be restricted or related, the more it is informationally enriched – the obverse of high entropy. Some people might even say that the structure or process is complex rather than informationally rich.

Complexity

Perhaps the most amazing miracle of our universe is that despite the seeming raw statistical power of entropy, complex organized structures like stars, galaxies, planets, life, and ultimately, brains capable of pondering such things came into being. Indeed, as David Christian points out in his book of big history, *Maps of Time,* "The endless waltz of chaos and complexity provides one of this book's unifying themes" (Christian, 2011, p511). Hence, to better understand big history, we also need to better understand complexity, or complex systems if you prefer. However, even at the Santa Fe Institute (SFI), which specializes in the study of complexity, there is no universally agreed upon definition for a "complex system." A few of the definitions offered by complexity science experts who were interviewed for SFI's 2016 online course, "An Introduction to Complexity," include (Note, some definitions below have been slightly condensed from their exact verbiage):

- "A system that has a very sophisticated internal causal architecture that stores and processes information." Jim Crutchfield, University of California, Davis.
- "A system that has interactions, nonlinear elements in it, and usually have scaling properties like power laws or fractal properties embedded in them." John Rundle, University of California, Davis
- "A system with a bunch of entities that may not start out being diverse, but end up being diverse, are connected in some way (usually a network structure or some spatial structure), and they get information through that network or spatial structure, but also sometimes get some global signals or information." (whew!) Scott Page, University of Michigan.
- "A system with many interacting components and the interactions between the components have nontrivial or nonlinear interactions and that leads to a system having unpredictable behavior." Stephanie Forrest, University of New Mexico
- "A system with a lot of interacting parts where something about the way those interacting parts behave is qualitatively different than the way they behave if you look at them individually." Doyne Farmer, University of Oxford
- "A system that contains enormous numbers of actors or agents that are interacting usually in a nonlinear fashion from which all kinds of multi-level behavior evolves so that there are emergent phenomena." Geoffrey West, Santa Fe Institute

What is common to all of these definitions is that they depend on describing various properties of a complex system, rather than a single, core characteristic. Indeed, noted big historian professor, Fred Spier, states in his book *Big History and the Future of Humanity*, "Because no generally accepted definition of 'complexity' appears to exist, I decided to tackle this problem by making an inventory of its major characteristics" He goes on to state, ". . . a regime is more complex when more and more varied connections and interactions take place among increasing numbers of more varied building blocks." (Spier, 2015, pp48-9). Resorting to a definition based on characteristics, is not unique to "complexity," because "life," and "civilization," complex systems in themselves, are also defined by their properties, e.g., "life" is something that is able to metabolize, reproduce, and evolve. In regards

to key characteristics of complex systems, at least two of the definitions from SFI faculty included the term "information." Nearly all the rest, including Spier's definition, include the terms "interacting" or "interactions," which is synonymous with the transfer of information from one entity to another - whether those entities are electrons exchanging photons with the nucleus, or the brain's hypothalamus interacting with the pituitary gland, which interacts with various other glands of the body. In other words, all of these definitions explicitly or implicitly include "information**"** as a key characteristic of complexity.

Note that none of these definitions included any mention of increased energy flow density through a system as an essential property of complexity, which Eric Chaisson convincingly demonstrated in his oft-cited book, *Cosmic Evolution* (Chaisson, 2001, p13). The exception is Spier, who does later discuss Chaisson's observation on this aspect of complexity (Spier, 2015 pp 53-64). Of course, the absence of "energy flow density" in those interviewed by SFI does not mean that Chaisson is amiss in noting and analyzing this phenomenon. It is more likely that he has discovered and quantified a unique and laudable insight into one of complexities key characteristics. Nevertheless, Chaisson also states that complexity can also be operationally defined "as a measure of the information needed to describe a system's structure and function." (Chaisson, 2001, p13). Hence, it is apparent that there is widespread consensus that information is a defining characteristic of complexity, even if it is often guised as "interactions." I will also assert that while energy flows are necessary for complexity to occur, it is not by itself sufficient. Information is also necessary and just as fundamental, if not more so. Consider: regardless of how finely tuned or how much energy is made to flow through the mended corpses that made up Frankenstein, the monster will never come to life. Too many proteins have denatured, the blood has clotted, neurons have withered, and too many cell membranes have lost their integrity. In short,

the many critical relationships or informational content of the body have been lost to entropy, and reanimation is not even remotely possible.

Therefore, it is the interwoven dance of at least the three fundamental ingredients of the universe: mass/energy, fields of force, and information that makes complexity possible over the course and stage of space-time. It might be that there is yet some other ingredient(s) that eventually made complex structures and processes like life and minds possible. The origins of these ultimate expressions of information and complexity have yet to be fully satisfactorily explained although complexity science is especially working hard to understand the origins and aspects of complex systems. While acknowledging that I am not giving complexity science the attention it deserves, I propose that we nevertheless, go forward and look through an information-centric lens to examine at least a <u>few</u> of the phenomena that have transpired through big history.

The Big Bang – and then there was "1"

Claude Shannon is credited with, along with many other aspects of information theory, determining that the most basic unit of information is a "bit" that is often represented by a binary digit -- a "0'" or a "1." A binary digit in turn represents any dichotomy such as "yes" or "no," "black" or "white," and even "existent" or "nonexistent." By analogy, we could state that at the instant of the Big Bang about 13.8 billion years ago, the cosmos went from a "0" to a "1" – John Wheeler's ultimate "It from bit." Still, the amount of information of the universe at 10^{-43} seconds was "0" because the logarithm of 1 in any base is that value. Intuitionally this makes some sense because the initial completely undifferentiated nascent universe was also indeterminably small, and indeterminably hot (Fewster, 2016, p35). In fact, it was so hot that there weren't even any "particles" to form relationships and, hence, informational content.

According to current cosmological theories, the fundamental forces of nature, gravity, strong nuclear

force, electroweak force, and then the 13 various fundamental particles of the standard model like quarks, photons, electrons, etc., all "precipitate" from the early roiling universe between 10^{-42} seconds and 10^{-6} seconds (Fewster, 2016, p34-5). With each new "ingredient" to the universe, the informational content would seem to increase by some vast new quantity. In fact, physicists estimate about 10^{80} fundamental particles exist in the observable universe (Seife, 2007 p65). However, the estimated informational content of the universe is calculated to be somewhere between 10^{90} and 10^{120} bits because you must also include other parameters like the particles' velocities, spin, mass, etc. (Schumacher, 2015a, p287-8).

An early example of interactions creating information is when hydrogen (~75%), helium (~25%), tiny amounts of deuterium (.01%) and even less lithium nuclei formed by about 3 minutes after the Big Bang (http://w.astro.berkeley.edu/~mwhite/darkmatter/bbn.html). It is worthwhile noting that physicists are not in total agreement whether information can be technically created or destroyed. However, at least new "kinds" of information or new relationships occur over time. The quarks and gluons are now interacting in novel ways to comprise protons, neutrons, and combinations of them to form atomic nuclei. It is also worthwhile noting that the information increase caused by the formation of these components is not predicted by "H = -k log M," because this formula works only if there were 2 different components that occurred with equal probability. The more general formula for the information content for an event that occurs with a different probability than a flip of a coin, i.e. other than a 50:50 chance, is "H = -k log $1/p(x)$" where "$p(x)$" is the probability of event "x" occurring (Stone, 2015, p36). This variation of informational quantity is also sometimes referred to as the "surprise" and often abbreviated as "$s(x)$" rather than "H." To rephrase as Shannon might have stated: "The greater the surprise of a message, or

the less likely it is to occur, the greater it reduces informational uncertainty." In the parlance that I have been proposing, improbable information is also very "novel."

This assertion is nicely illustrated by noting that the observed ratios of deuterium (one proton and one neutron in its nucleus) and helium nuclei that astrophysicists observe in interstellar space are the same as those that they calculated would have formed during a brief period early after the Big Bang: 0.0001 deuterium and 0.23 helium, the remainder being hydrogen and a tiny bit of lithium. (http://w.astro.berkeley.edu/~mwhite/darkmatter/bbn.html). The high "surprise" of that small presence of deuterium ($\log_2 1/.0001 = 13.29$ bits) helped to convince many cosmologists that the proposed Big Bang model was correct, i.e. the rare occurrence of deuterium and its accompanying high informational value was strong evidence that their theories were correct. Adding the informational "surprise" of the correct amount of predicted helium further substantiated the Big Bang theory ($\log_2 1/.23 = 2.12$ bits), but not as much as detecting the predicted small amount of deuterium – at least from a purely informational theory perspective. (Note: for simplicity, the "-k" value was ignored as it is in many sources because it does not change the conclusions.)

The End of the Dark Age – Information Gets to Travel!

If the proportion of deuterium to "regular" hydrogen and helium nuclei is disproportional, consider photons. Photons, one of the fundamental particles of the universe, outnumbered quarks and other particles by a factor of at least 1 *billion* to one after the annihilation of particles and anti-particles ceased about one second after the Big Bang (Christian, 2011, p25). Those photons and their distribution are represented by the "cosmic microwave background" (CMB) which was famously discovered by Penzias and Wilson in 1965, and later mapped by the COBE and WMAP satellites. The

CMB does not represent the "same" photons from the first moments after the Big Bang, but rather those that began to scatter 380,000 years later when the universe had cooled sufficiently to allow the freely roaming electrons to be captured by the nuclei to form complete atoms (Fewster, 2015 p34). With the capture of free electrons, the photons were not continually being absorbed and reemitted after traveling short distances. Now the photons could travel unimpeded not only through space, but through time as well such that some of them ended up on the radio antenna of Penzias and Wilson, or the detectors on board the COBE and WMAP satellites. This event 380,000 years after the Big Bang is called the end of the cosmic "dark age" (Spier, 2015, p90).

The photons that travelled through time and space for ~13.8 billion years to land on our detectors, not only gave us information about that event, they also demonstrate another feature of information – the fastest at which it can travel. Due to the constraints of known physics, the fastest that *anything* can travel is the speed of light through vacuum, about 300,000 kilometers per second. Another way to think of this fact is that nothing in one part of the universe can affect or relate to another part of the universe sooner than it takes a photon to travel that distance. As a side note, it is interesting that especially in the past, historians referred to time periods when there was paucity of or a decrease in information as a "dark age" such as the Greek or Medieval Dark Ages. In other words, we intuitively, or coincidentally at least, associate light with information.

Also, the estimation that only 1 in a billion elementary particles are *not* the nearly evenly scattered photons of the CMB indicates that a great amount of the universe's entropy was "created" right after the Big Bang. Besides being the obverse of information, entropy is also a measure of energy that is not available to do work – and energy to do work is necessary to make complex systems as Chaisson rigorously pointed out. Despite such a large dissipation of energy in the first moments after the

Big Bang, there were still enough energy differentials and concomitant low enough entropy to drive the creation of complex entities from stars to rain forests.

Increasing Complexity - the Gift of Information, Energy Flows & Time

Over the next approximately 10 billion years, the fundamental forces and particles created after the Big Bang with the added assistance of dark matter (whatever that is) went on to form stars ca. 13.6 billion years ago (BYA), super nova ca. 13.5_{BYA}, galaxies 13.4_{BYA} (Fewster, 2015, pp44-5), and at some point in time, planets . Note that the sizes of these structures ranging from small asteroids to the eventual galaxy superclusters are vastly larger organized structures than the preceding atoms or nuclei of the primordial gas cloud. Gravity was the instrumental force for creating these much larger and more complex entities. At first, glance, this increase in order would seem to be a contradiction to the second law of thermodynamics which indicates that entropy will remain the same or increase with the passage of time. Recall, however, that *local* entropy can decrease as long as the universe's (the ultimate "isolated system") overall entropy increases.

Restated in terms of information, stars and galaxies require much more information to describe their structure and processes than would a similar amount of an amorphous gas cloud – like a nebula. Although, it would take a great deal of information to describe the relative positions, directions interactions, speeds of travel, compositions, etc. of each of the nebula's particles, it would require even more information to describe those same parameters, plus their ordering, more varied density, new interactions (informational or relational changes), and newly created particles, like carbon to name a few. This kind of analysis, although to a much more profound depth, led Erwin Schrödinger (1887-1961) of quantum mechanics fame, to call this process of localized ordering and informational increase as "negative entropy" in his book *What is Life?* (Schrödinger, 1967, p71).

The phrase was later shortened to "negentropy" by another physicist Léon Brillouin (1889-1969), in part to avoid the word "negative" with its associated connotations (http://www.informationphilosopher.com/solutions/scientists/brillouin/).

As briefly alluded to above and importantly for the future of complexity, stars increase the variety of nuclei, and eventually atoms, by forging elements up to iron in their cores, and elements up to uranium when they explode as supernova (Fewster, 2015, p63). Although the vast majority of elemental atoms in the universe are still hydrogen (about 90-92%) and helium (about 8-10%) (DeGrasse Tyson, 2004, p. 72), the remaining ~1% of the other approximately 90 natural elements are critical for the eventual creation of evermore complex structures and processes. The addition of the extra elements allows for a tremendous increase in the number of new possible relationships as the elements can combine with each other in innumerable ways – especially carbon which can form 10 million or more combinations with itself and other elements. Carbon also happens to be the highest element that can be forged in a star the size of our sun.

Life – Complexity takes Information Really Seriously

Although there are admittedly remaining profoundly important mysteries, we can satisfactorily explain much about the structures and processes of the universe with the known constituents and forces of physics and chemistry. Indeed, especially to a non-physicist like myself, it seems a staggering feat of accomplishment that scientists can accurately predict the magnetic strength of an electron to the 12th decimal place (Pollock, 2001, p121), postulate what happened at 10^{-30} second after the beginning of the Big Bang, and so on. Nevertheless, another complex system that still defies satisfactory explanation confronts us every day we look in the mirror, play with our pet, or even squash a mosquito – life. Also, if you are still skeptical about information playing a

real role in complexity, life is the phenomenon where Chaisson's energy flow density via metabolism more obviously entwines with Shannon's informational flow via reproduction, evolution, and other functions of DNA and RNA.

But how did energy and informational flows come to be so complex in themselves, while also complexly entwined? Although hypotheses abound as to how life generally came to *be* shortly after the Earth cooled sufficiently for it to exist, no current theory satisfactorily entirely predicts how this complex phenomenon not only originated, but persisted and spread to a truly remarkable degree. David Christian said it well, ". . . but at the biological level of complexity, new rules appear as well. Living organisms operate according to distinctive and more open-ended rules of change, which are superimposed on the simpler and more deterministic rules of physics and chemistry." Also, "So to understand living things, we need a new paradigm, one that takes us beyond the rules of nuclear physics, chemistry, or geology and into the realm of biology" (Christian, 2011, p81). Professor Christian also seems to give primacy to high energy flows as the defining characteristic of complex life: "The rules of biology are made possible by the high degree of precision with which living organisms reproduce. Handling large energy flows are such a delicate task that it requires extremely precise mechanisms; the rule book for creating and re-creating such structures has to be complex, exact, and accurate" (Christian, 2011, p81). Admittedly, metabolism is one of the defining features of life and it has all the features that Christian mentions. However, I would assert that the information needed to realize the complex mechanisms of metabolism, as well as reproduction and evolution is co-equal to energy flows, if not paramount.

Admittedly, it is unlikely that complex interactions or information flows of complex systems would be possible without the other – high, finely tuned energy flows. They are tied together like a Gordian

knot. An organism perishes when either energy flows are insufficient (e.g., insufficient food, cyanide poisoning), or information flows are disrupted (e.g., neurodegenerative disorders, proteins denatured by high temperatures), or both (e.g., respiratory or circulatory failure).   Nevertheless, aging and death itself is inevitable, not primarily due to failing energy flows, but because of the inexorable march of entropy which causes complex relationships to steadily degrade over time: the skin wrinkles and sags, the hair greys, bones become brittle, and, yes, the heart's output declines as well, but typically due to various changes in its tissues.

Yet, on the other hand one of the most profound miracles of life is that it can also repeatedly and faithfully renew its information virtually unchanged via reproduction despite entropy, even over billions of years as in the case of bacteria or archaea. A miracle of similar magnitude is that life has also diversified its informational content into literally 100's of millions of species over time, and with even greater degrees of complexity via evolution. That is, the information of life can both replicate itself accurately, while also occasionally varying its replication such that it has also increased its depth and breadth over time. In the final analysis, it would seem that life especially exemplifies that energy flow is the hand maiden of informational flow. If still in doubt, consider viruses – packets of information that hijack a "true" life form's metabolism to reproduce itself. Note that there is no known equivalent entity constituted primarily of an energy structure like a mitochondrion that hijacks a true life form's informational contents to reproduce.

Admittedly, one of the advantages of energy flows in complex systems is that it is more readily calculably quantified – and scientists are often understandably enamored with mathematics and its quantitative predictions. Even from an informational centric viewpoint this love affair makes sense: math is but the very precise pronouncement of how relationships work and often makes these

pronouncements much more scientifically testable. While energy flow densities can sometimes be precisely predicted and stated in mathematical terms, the mathematics associated with information theory often predict limits rather than exact quantities of informational content, change, effects of noise (informational interference), and other parameters.

One example of information theory's ability to predict limitations is in regards to determining the minimum number of symbols or codes needed to convey a message. To illustrate, DNA is fundamentally a set of codes that directs the reproduction and many of the functions of living organisms. One very common type of DNA based instruction is how to sequence any of the up to twenty amino acids available to make particular proteins. To determine the minimum number of bits required to represent or code for these twenty amino acids, you must also have a code for the command to "start" and "stop" making the protein. To determine the minimum number of needed codes, you take the $\log_2$ of twenty-two, which equals 4.46 bits. Pragmatically, you can't use 0.46 of a symbol to represent a bit, so DNA has to round up to at least five bits to represent the twenty-two necessary codes. As it turns out, DNA uses four different nucleotides abbreviated as A (adenine), G (guanine), C (cytosine), and T (thymine) in sets of three to comprise those codes. For example, the DNA code "GGU" represents the amino acid glycine. The possible number of permutations of four nucleotides in sets of three is equal to 4^3 or sixty-four. The $\log_2$ of sixty-four is six bits. Six bits is greater than five bits, which means that the DNA coding for protein synthesis satisfies the rule for the *minimum* number of codes necessary for a message, which in this case is a completed protein.

If nature was solely concerned with efficiency, it might have instead chosen to use codes comprised of five nucleotides in sets of two, which gives you 2^5 possible codes or exactly five bits. However, life has to worry about more than efficiency. As

Shannon would state, life also has to contend with the "noise" in the signal channel. In this case the cellular cytoplasm for prokaryotes (bacteria and archae) or the cell's nucleus for eukaryotes (other life forms) is the communication channel, between DNA and the environment. Reactive chemicals, radiation, and thermal motion, to name a few factors, are some of the sources of noise that can cause an unintended change in DNA's code sequences. Having six bits of code rather than the minimum five bits allows for increased redundancy in the code so that not all noise induced changes of the DNA (i.e. mutations) lead to potentially harmful alterations in protein synthesis. Glycine, as an example, is symbolized by GGG, GGA, GGT, as well as GGU. Similarly, other amino acids and the "stop-making-the-protein" codes are also represented with several similar sequences. Hence, an inadvertent change in one nucleotide does not always result in the dysfunction or even death of an organism from an altered protein.

Information theory can also provide other insights into life such as why the exchange of DNA via sex might have evolutionary advantages over relying primarily on mutations as in asexual organisms (Stone, 2015, pp.188-193), and the upper limits of mutations that early precursor molecules for life would be able to tolerate without failing to reproduce – the so called Eigen error catastrophe. (Schumacher, 2015, pp156-8). However, at this time at least, we must concede that there is not a quick, clear correlation between the number of genes, which are rough "units" of DNA information, and the complexity of an organism. A recent study, for example, determined that the human genome contains less than 20,000 genes, which is far fewer than that of a water flea which has 31,000 genes (https://www.popsci.com/article/science/humans-may-have-fewer-genes-worms). Even though this apparent paradox might be explained by other factors such as it is also *how* genes are controlled by non-protein coding regions of DNA that defines an organism's complexity, an easy and seemingly

obvious metric for measuring the complexity of an organism is not as readily available as Chaisson's energy flow densities – at least at this time.

The Brain – The Ultimate in Complexity and Information Processing

If complexity is best measured by energy flow density, then the brains of humans and other "higher" animals surely qualifies as among the most complex systems based on that metric alone. As calculated by Professor Chaisson, the brain uses about 150,000 ergs/sec/gm whereas the body overall uses about 20,000 ergs/sec/gm, or about 7.5 times the body's average (Chaisson, 2001, pp138-9). However, the design and purpose of the brain is not to simply expend energy, but rather to access, process, store, and transmit information. Again, energy flow is serving the needs of information flow for both the design and function of a complex system.

The biological-based neurosciences and information sciences have gone a long way to describe many of the secrets of how the brain works. We know much about what area(s) of the brain serve which functions, how stimulated neurons transmit electrical potentials down their length to cause the release of varied chemicals at its far ends to pass on a signal to the next neuron, that the brain can only process and retain about 2.5 bits of one type of sensory information at any given moment (Schumacher, 2015a, p171), and so much more. In fact, there is a sophisticated level of research called "computational neuroscience" dedicated to applying information theory to the workings of neurons and large neural systems (Stone, 2015, p195).

Nevertheless, when you consider higher functions of a brain as advanced as a human's, we still have a "black box" of complexity from which emerges incredibly surprising phenomena like self-awareness, emotions and other subjective experiences, future anticipation, past reflection, and abstract problem solving to name a few. If you were some disembodied, detached super-physicist

present at the Big Bang, would you be able to predict that the various fundamental particles and forces of nature could relate together in just the right way to eventually create such strange phenomena? Furthermore, while these higher functions remain a fundamentally deep mystery, it is much more likely a manifestation of informational processing, integration, and feedback loops, than a result of finely tuned energy flows, even if the latter is a prerequisite.

Humans Take Information and Complexity for a Ride

One of the most important "tricks" of the human brain is its advanced ability to turn syntactical information into semantical information. To reiterate from earlier, syntactical information is the raw ordered structures and processes "out there" in the world, whether it was created by a natural process or another "agent." The human ability to apprehend syntactical information, process it via the brain, give it rich semantical content, and then communicate that information to others was possibly the single greatest set of related abilities that led to us eventually dominating the planet – for better and for worse. As mentioned earlier, semantics began when the first organism detected something within itself or in its surrounding environment, and responded to it in some manner, e.g., it sensed a depletion of nutrients and slowed its metabolism. At this simple level, you might feel that it is a stretch to claim that the syntactical information it gained from its environment caused a simple organism to "purposely" slow energy expenditure when the information flow was likely a fairly direct, even if long, sequence of chemical reactions. However, as life diversified, some forms increased the complexity associated with detecting, processing, responding, and eventually becoming aware of at least some information to which it was exposed. Subsequently, the meaning of semantics becomes ever more meaningful in itself. Somewhere during evolution, at least by the time a central nervous system

develops, it becomes ever more difficult to trace a clear path of syntactical relationships from sensory input to some output that doesn't beg us to identify other phenomena, like awareness, anticipation, memory, etc. Information becomes not just a series of morphing relationships, but morphs itself into an agent for which information of its external and internal environment carries ever deeper, and dare I say, more complex meaning.

Humans seem to be the epitome of conscious agents and are able to give semantical content to even the simplest syntactical sensory data. Religious symbols, national flags, and the musical notes of "Taps," are but a few examples of humans communicating abstract, rich information from one to another via fairly simple symbols or signals. The beginning of this "symbolic thinking" began for certain by the time of the earliest cave paintings around 38,000 b.c.e. (https://www.nytimes.com/2014/10/09/science/ancient-indonesian-find-may-rival-oldest-known-cave-art.html). It is possible, however, that it began as early as about 80,000 b.c.e. as suggested by the presence of ochre, likely used for body decorative purposes, that was discovered in a cave in South Africa (http://www.nytimes.com/2002/02/26/science/when-humans-became-human.html).

The earliest evidence of symbolic communication is visual because pigments on walls, or materials like ochre in protected areas were able to survive the passage of considerable time. However, humans have historically used complicated and ephemeral sounds to communicate most of its information to others, and likely did so at least as early as our use of visual symbols. Despite its transient nature, the choice of using varying sound waves to communicate makes sense from a physics and environmental perspective. First, sound travels quickly, about 1,000 feet or 330 meters per second. Another option might be odors, but the speed of travel would be limited by wind speed and thermal motion. Another option

would be the fastest possible option, light. However, light waves are easily reduced or entirely blocked by common things in the terrestrial environment like plants, rocks, and hills. Also, because we don't have an organ or tissue that emits light, like a firefly or angler fish, communication by this modality doesn't work in the dark of night. Touch, another sophisticated sense, is used for some communication, but is obviously limited in extent by one's reach. Therefore, the speed and transmissibility of sound make it a good choice for warning, finding, and generally informing others.

The human body is also designed to emit a much larger variety of sounds than light (e.g., skin color change) or odors (e.g., pheromones) and, therefore, can communicate a much greater variety of messages which can even be nuanced by inflection, musicality, loudness, sound order, and other variables. Finally, we can change and exchange the utterance of sounds much faster than we can change colors or odors. In the parlance of information, the ordered utterance of changing sound waves allowed for the faster and omnidirectional communication of bits of information through space with less interference from background noise. It also allowed for a greater diversity of bits of information to be quickly communicated. Finally, although various species communicate to each other by changes in light waves or patterns, odors, touch, various sounds, and sometimes by even other means (e.g., bioelectrical fields), it was the progressive evolution of an ever richer use of sounds that would eventually become "language." The semantic richness of language in turn made us capable of a much greater range and depth of "collective learning" compared to other life forms (Christian, 2011, p146-7).

But still, there is that ephemeral problem. While sound travels well through a reasonable range of space, it does not travel well through time. Oral traditions do mitigate this problem, but rely on the memories of a chain of individuals which can introduce a significant amount of noise so that

the original information becomes corrupted, as it commonly does with social gossip. Humans developed techniques to reduce the noise of memory through the use of meter, rhymes, repetition, musicality, and other means to better communicate lengthy bits of information, like the Homeric epics, to later generations (Gleick, 2011, pp34-5). Still, having an informational medium as rich as vocal sounds, but as long lasting as visual signs would potentially convey much more information, with less alterations from memory noise, to more people over longer periods of time. In other words, it would be nice to have a way for the collective learning from one generation to be more accurately and extensively passed onto the future ones. Restated as the core central theme of this paper, it would be advantageous for humans to be able to more permanently, richly, extensively, and reliably communicate learned relational data to others over greater distances of space as well as time. Enter the written language.

Creating a rich written language is a rare and apparently difficult achievement. It was created from "scratch" only three, possibly four times in human history: by the Sumerian, Chinese, Mayan, and possibly the Egyptians. Whether Egyptians developed writing independent of Sumeria is a matter of contention among historians (Parker, 1986, pp50-1, 262). As you can tell from the names of its originators, the development of writing apparently requires a "high" civilization as a cultural milieu. Civilizations in turn are dependent on the development of agriculture. Writing or even other forms of semantically rich visual communication, like the Inca knotted ropes (Quipus) never began in hunter-gatherer or pastoral nomad societies. This sequence of events nicely illustrates the interplay that occurs between information and energy flows for promoting the development of complex systems. To wit, agriculture's primary role is to increase the availability, reliability, and locality of energy flows from the sun to humans via the cultivation of plants, and the utilization and consumption of domesticated

animals. This increase in energy flow, in turn made possible the development of civilizations, which used this energy to increase its relational or informational complexity via a more divided and hierarchal social structure, increasingly sophisticated material goods, and grand architecture, to name a few of its salient features. Civilization in turn found it necessary to develop a better way to record information for pragmatic purposes like inventories, taxation, the coordination of work or war projects, as well as for spiritual, aesthetic, and other reasons.

Writing went through substantial improvements over its subsequent history in regards to its cost, portability, decreased errors in reproduction, ease of manufacture and access. Think clay tablets versus papyrus or paper, scrolls versus codex, and writing advancements like the invention of the alphabet, word spacing, Carolingian miniscule, and punctuation. Perhaps the most important improvement responsible for propelling the next great leap in human social complexity was the invention of the printing press by the Chinese in the first century, c.e. (Fewster, 2016, p267) which was then improved further by the European, Johannes Gutenberg, around 1440 c.e. The improved the printing press together with the more printing-press-friendly Western alphabet, subsequently increased collective learning by several magnitudes for all the reasons given above. Once again the invention of the Gutenberg printing press and the subsequent sequence of major events help to illustrate the interplay between energy and informational flows that can occur and result in increased complexity.

First, the printing press fundamentally made information flows through societies much more efficient, and thereby pervasive. Arguably, the first major impact from the printing press was its effect on the Catholic religion in Europe. The widespread printing of both diversified religious views and the Bible itself into its traditional Latin as well as vernacular languages made it essentially impossible for the Catholic Church to monopolize

Biblical information as it had before. Subsequently, it could not fully quell the informational variants of the "word of God," (heresies) as it had with earlier movements like the Albigensians, Gnostics, Monophysites, and others. This spread of diversified religious information in Europe certainly did add new complexities to the political and spiritual structures and processes of the continent, not to mention the catastrophic Thirty Years' War (1618 - 1648). However, it would likely be hard to argue that the increased European religious diversification that was promoted by the printing press created any novel social complexities that weren't already present in other locales, even within Europe. For example, the Iberian peninsula had long been religiously diversified with Muslims, Christians, and Jews living together under the Umayyads. the Indian subcontinent in particular was already host to an even more diversified mixture of Hindus, Muslims, Jains, Buddhists, and the early Sikhs. Complexity changed at a much greater rate, however, when books helped to both precipitate and more quickly disseminate two of the major revolutions in human history: the scientific and industrial revolutions.

The scientific revolution was informationally driven. Although, a more rigorous scientific way of understanding the world had earlier beginnings, like Copernicus' (1473-1543) publication of *[De revolutionibus orbium coelestium](#)* in 1543, it arguably began in earnest with the empirical studies of Galileo (1564-1642) and the printing of Francis Bacon's (1561-1626) *Novum Organum Scientiarium*. Both of these events occurred in the first quarter of the 1600's and modern science gained steady momentum from that time forward. Importantly, Galileo's work and Bacon's treatise demonstrated and carefully explained, respectively, a more rigorous way to determine if a rational proposal about how the universe works does indeed coincide with reality empirically. In raw informational terms, does 10111 "+" 01001 "=" 1011101001 *as predicted* or not? (Note: this example is purely

fictional and oversimplified, but simply meant to illustrate a general point.) Once the works of Galileo and especially Isaac Newton (1642-1727) proved the success of this approach, major shifts in informational authority (e.g., church versus scientific community), the rate of progress, and institutional changes began to accelerate. Information flows were also augmented by extending our senses, at first visually with the inventions of the telescope and later the microscope. Later inventions not only augmented the information we gain from our existing senses like hearing, sense of time and direction, but also extended our ability to gain information from phenomena that are entirely removed from our senses, e.g., radio waves, magnetism, radioactivity, x-rays.

Another important "revolution" that must be mentioned, even if only briefly because of its huge impact, is the "Columbian exchange." That is, the beginning of the first truly global exchange of information, people, and materials that began with Columbus' voyage in 1492. The exchange rates, variety of items, and trade distances would quickly eclipse those of earlier trade networks like the "silk road." To some extent even energy flows increased somewhat as calorie rich crops like the potato and sugar cane were cultivated in new lands.

The Modern Age – Information & Energy Positive Feedback Loops

Still, the scientific and Columbian exchange revolutions did not appreciably change the day-to-day lives of the great majority of people in the Old World (Christian, 2008, p220). While a myriad number of reasons for a greatly accelerated rate of change can be forwarded, the key reason is the onset of the industrial revolution in the mid-18th century. England, the first industrialized country, added a substantial increase in energy flow rates with the invention of the steam engine and many other inventions that harnessed the energy of its readily accessible coal.

With the already extant printing press, a more widely educated population, global exchange networks, and scientific method, a positive feedback loop was created where an improvement in one invention led to a cascading fount of other improved and diverse inventions, which led to even more innovations, which rapidly spread to other parts of the globe where differences in culture, resources, or simple intellectual talent could add to innovations further. Science and an educated populace were key players even early on in this feedback loop and the industrial revolution quickly morphed into the technological revolution – industry combined with science if you will. This dovetailing of industry and science began early in the industrial revolution when people like Sadi Carnot (1796-1832) and Rudolf Clausius (1882-1888) tried to understand how to make steam engines work more efficiently, and if it was even possible for *all* the energy put into a steam engine to be transformed into work. Their intellectual efforts in turn gave birth to a foundational area of physics, thermodynamics. As discussed earlier, later pioneers like Ludwig Boltzmann and the American physicist, Willard Gibbs (1839-1903) developed an even deeper understanding of thermodynamics and its core tenets like entropy. This led eventually to Claude Shannon, John von Neumann (1903-1957), Rolf Landauer, and other 20th century thinkers to finding the link between thermodynamics and information. Now, we have come full circle back to information theory.

Of course, to this day, the interplay between energy flows and informational flows continues to propel human social changes and complexity at an astounding rate – for both better and worse. On the side of "better," humanity has not seen a Malthusian crisis of population crashes via mass starvation or epidemics due to advancements like inexpensive crop fertilizers, clean water supply, and vaccines to name a few. Even the Spanish flu, the worst epidemic of the modern era killed "only" up to 3.3% of the population (http://www.history.com/topics/1918-

flu-pandemic) versus the black death which may have killed up to 33% of Europe's population in the 14th century (http://www.history.com/topics/black-death).   However, it is also evident that our increasing population and social complexity, with its extraordinary demands on our planet's limited resources, comes at the cost of damaging another ancient, unique, invaluable, and incredibly complex system - the Earth's biosphere.

Information and Complexity – A Conclusion

The foregoing discussion is a brief introduction of why information is an inseparable, integral aspect of complexity. Of necessity given the space allowed, the review has been both superficial in depth, and incomplete in scope. Indeed, information science has made contributions to many disciplines from computer science, to economics, to sociology. Much more mathematics and other insights can also be offered to describe or predict various phenomena from the amount of information believed to reside on the surface of black holes, to how much information the brain is capable of storing. Complexity science also offers more profound insights and math for examining and predicting features of complex systems, and has even discovered other surprising sources of ontological indeterminism (Mitchell, 2009, p33). Quantum mechanics and thermodynamics are not the only disciplines to discover that the universe is ultimately statistical rather than deterministic – the "clockwork universe" was a mirage.

Information theory and complexity science will consequently be a rich fount from which big historians can better analyze and understand countless events, and processes that have occurred over time. Likewise, information and complexity scientists will find big history to be a rich source on which to apply their insights on this inherently rich and cohesive multi-disciplinary project. After all, even though the 20th century will be remembered in part as the time when relativity, quantum mechanics, and information theory were all discovered, it is still likely that our contemporary age will continue to be remembered *not* as "the "relativity period," or "the quantum era," but instead as "the information age."

––––––––––––––––

Side bar: A brief overview of logarithms

If you are like me, it was decades since I had done any mathematics involving logarithms. Fortunately, basic logarithmic mathematics is relatively straight forward, and although not absolutely essential to understand the basic concepts behind both the second law of thermodynamics and information theory, it is quite helpful.

A logarithm is expressed in a "base" that is some number greater than 1. One of the most common logarithms (log for short) is expressed in base 10 and formally shown as $\log_{10}$. Many times, however, the subscript is left off and is simply shown as "log." The log of a number is what that number would be if 10 was increased by some exponent. For example $10^3 = 1000$, therefore, log 1000 = 3. $10^6 = 1,000,000$ and, therefore, the log of 1,000,000 is 6, and so on. The log of some number between 1 thousand and 1 million would similarly be between 3 and 6. A calculator can show you for example that log 5700 = 3.7559. In other words, $10^{3.7559} = 5,700$.

One of the obvious advantages of logs is that it makes it easier to express very large numbers. This feature is useful in thermodynamics where a vast number of microstates are possible for a system, or in information theory where a similarly large amount of data is involved. An especially importantly feature of logs for information theory is that if you combine the logs of information of two sources, the logs are additive rather than multiplicative. To illustrate this importance, imagine that you have two books of the same size that cover two entirely different topics.

If you combine the different number of possible messages from both books, represented as B1 and B2, you would have B1 x B2 = $B(1 + 2)^2$ possible informational content. However, your intuition tells you that you would not square the amount of information that you gain by reading two separate books, but instead it should be doubled at most. Logs solve this problem by being additive for the increase in informational content rather than multiplicative. In this example, using the rules of log: log (B1 x B2) = Log B1 + Log B2.

Other important log rules:

$Log_b(x^n) = n\ Log_b\ x$

$Log_b(x/y) = Log_b\ x - Log_b\ y$

$Log_b(1/x) = -\ Log_b\ x$

$Log_b\ 1 = 0$ (regardless of which base is used)

Knowing these rules is important if you decide to read some source on information theory or thermodynamics, because different authors will often use different appearing versions of the same equation (and sometimes with different letters to represent the same variable) – which can be confusing to say the least, e.g., most information science books use "H" to represent information content, while Chaisson's book *Cosmic Evolution*, uses "I.". It's also important to know that the log base used, whether it is 10, 2, "*e*," or some other value is arbitrary and doesn't fundamentally change the equation except for the value of an accompanying constant, often denoted as "k." (Note: "*e*" or "Euler's number" is an irrational number that mathematicians often use. When used as a log base, is called a "natural log" and often abbreviated as "ln.")

Because information theory's preferred numbering system is "binary," the log base used in information is typically "2." Therefore, because $2^1 = 2$, $2^2 = 4$, and so on, the log_2 of a number gives you the number of bits involved. For example, if you want to determine the minimum number of bits needed to communicate using only the upper case letters of the alphabet plus a space, you will need at least $log_2\ 27 = 4.75$ bits. Because you can't pragmatically have 0.75 of a bit

to actually use in practice, you will need a minimum of 5 bits to communicate this way, e.g., a = 00001, b = 00010, and so on. ASCII is a code commonly used in computer programming and has 7 bits to represent all the symbols on a standard Western keyboard. 2^7 =128 possible bit combinations, which allows all the symbols (a-z, A-Z, 0-9, #,@, etc.) on the keyboard to be represented by its own unique binary code. Hence, in information theory and computer science, the log base is considered to be a "2" as a default and is frequently not indicated in that literature.

––––––––––––––––

References

Bateson, G., 1972, *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology.* Chicago, University of Chicago Press

Chaisson, Eric J., 2001. *Cosmic Evolution, The Rise of Complexity in Nature*, Harvard University Press

Christian, David, 2011. *Maps of Time, an Introduction to Big History,* Berkely and Los Angeles, California, University of California Press

Christian, David, 2008. *Big History: The Big Bang, Life on Earth and the Rise of Humanity,* Chantilly, Virginia, The Teaching Company

Deacon, Terrence W., 2010. "What is Missing from Theories of Information?" In *Information and the Nature of Reality, From Physics to Metaphysics,* edited by Paul Davies and Niels Henrik Gregersen, New York, Cambridge University Press

DeGrasse Tyson, Neil and Goldsmith, Donald, 2004. *Origins, Fourteen Billion Years of Cosmic Evolution*, New York, W.W. Norton & Company, Inc.

Doyle, Bob. www.theinformationphilosopher.com

Fewster, Helen, et al, 2016. *Big History, Examines our Past, Explains our Present, Imagines our Future,* New York, DK Publishing

Floridi, Luciano, 2011. *The Philosophy of Information*, Oxford University Press

Gleick, James, 2011, *The Information, a History,*

a Theory, A Flood, New York, Vintage Books

Layzer, David, 1990. *Cosmogenesis, The Growth of Order in the Universe*, Oxford University Press

Layzer, David, 1970. "Cosmic Evolution and Thermodynamic Irreversibility," Pure and Applied Chemistry, v.22, p.464

Mitchell, Melanie, 2009. *Complexity, a Guided Tour,* Oxford University Press

Parker, Geoffrey, 1986. *The World, An Illustrated History*, New York, Harper & Row Publishers

Pollock, Steven, 2003. *Particle Physics for the Non-Physicist: A Tour of the Microcosmos Guidebook*, The Great Courses, Chantilly, VA

Schumacher, Benjamin, 2015. *The Science of Information, From Language to Black Holes*, Course Guidebook, Chantilly, VA: The Great Courses

Schumacher, Benjamin, 2015. *The Science of Information, From Language to Black Holes*, Lecture Transcripts, p562, Chantilly, VA: The Great Courses

Schrodinger, Erwin, 1967. *What is Life?*, Cambridge, U.K., Cambridge University Press

Seif, Charles, 2007. *Decoding the Universe, How the New Science of Information is Explaining Everything in the Cosmos, from our Brains to Black Holes*, London, England, Penguin Books

Shannon, C.E. 1948, "A Mathematical Theory of Communication," The Bell System Technical Journal, Vol. 27, pp 379-423, 623-656, American Telephone & Telegraph Co.,

Shannon, C.E. 1993, *Collected Papers,* ed. By N.J.A. Sloane and A.D. Wyner, New York, IEEE Press

Spier, Fred, 2015. *Big History and the Future of Humanity,* Chichester, West Sussex, UK, John Wiley & Sons, Ltd.

Stone, James V., 2015, *Information Theory, A Tutorial Introduction,* Lexington, KY, Sebtel Press

Wheeler, John A., 1990. "Information, physics, quantum: The search for links". In Zurek, Wojciech Hubert. Complexity, Entropy, and the Physics of Information. Redwood City, California: Addison-Wesley