A Proposed New Information Equation and its Application

"Brookfield Information"

Abstract:
"Information" has been understood mathematically for a long time by Claude Shannon's famous equation -- wherein one unit of information (i) is given by the negative log of the probability of the event (typically to base two). While this approach has been very successful technologically and has ushered in the "information age," not all is well with this definition theoretically and theoretical problems stubbornly persist in the effort to unite information theory, thermodynamics and black hole dynamics. The reason, as I see it, is Shannon's failure to mathematically differentiate *information* from *gibberish* wherein information is measured by its *improbability* and yet gibberish configurations are equally improbable. The equation I propose here makes this distinction clear while still supporting the technological validity of Shannon's information equation as a useful shorthand. I also provide some examples of my equation in application.

The proposed equation in question is as follows;

"Brookfield Information" $(Br\{i\}) = \sqrt{(-\log_2\ \{e\}P)(-\log_2\ \{c\}P')}$ where the "P" is the improbability of the informational event. "Brookfield Information" is therefore equal to the square root of the product of Shannon "Information" ($-\log_2$ {e}P) and the Correlational "Information" ($-\log_2$ {c}P′).

As I see it, another name for "Brookfield Information" would be "real information" or "actual information." This is because I see the correlated nature of information – information's "aboutness" -- as fundamental to information's two-dimensional nature. If there is no correlation between the email that you sent your friend and the email your friend receives, then the information for your friend has been lost. The mere receipt of the uncorrelated bits will not do. The second component ($-\log_2${c}P) therefore cannot be removed without a destructive over-reduction. This second component/dimension when coupled with the first produces a perfect orthogonality (when mapped XY) and corresponds at a microscopic level to William Dembski's "independently given pattern." Each correlational vector is a simple, but nonetheless "independent non-random pattern." Together these correlations summate to a larger detachable pattern (of co-extended parallel lines) that has no causal antecedent in the material world. Material forces alone make no requirement that these lines/vectors/correlations must be parallel, nor that they even exist. I am suggesting therefore that the *materially transcendent correlated bit* is the fundamental unit of information and that the more *transcendent correlated bits,* the more a design inference is warranted. Information is always information *about* something and "correlational information" is a mathematical formalization of "aboutness."

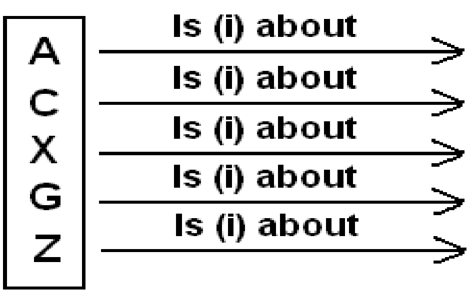

fig #1 Specified Correlational Order

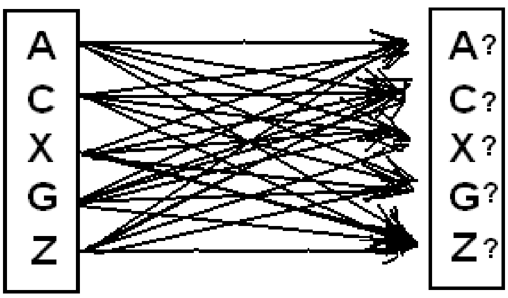

fig #2. Unspecified Uncorrelated Randomness

Example One: The Random Coin Toss.

Flip a coin 100 times at random, write the results down, and you will have experienced a highly *improbable* outcome. Using Shannon's equation alone, this high improbability is translated into a high "information" reading. The coin pattern however, being uncorrelated with anything beyond itself is devoid of what I call "real information."

It is worth noting that there exists a sister theory of information to Shannon's theory. In Algorithmic Information Theory (Kolmogorov & Chaitin) a random coin toss typically exhibits high K-complexity and subsequent high algorithmic "information." Algorithmic information theory thus suffers from the same problem as Shannon "Information" regarding randomness and information – gibberish and Shakespeare…

Example Two: Monkeys Typing "shakespeare."

In this case we have what initially appears to be correlations and subsequent "real information." The output appears to correlate with the rules of spelling, grammar, and plot development. Appearances however, can be deceiving and upon further examination this output is fleeting and appears only at an extremely rare frequency, fully consistent with a uniform probability distribution (randomness). The "shakespeare" pattern is actually necessary (along with all the other available patterns) to complete the uniform distribution – IE to make it uniform/random/"fair" – just as the three-lemon "jackpot" is necessary to make a slot machine uniform/random/"fair." The rules of spelling, grammar and plot development have no effect on the output and have been mistakenly projected (from outside) onto this one rare sequence. Being uncorrelated with anything beyond itself, monkey "shakespeare" is devoid of real information.

Without any targeting being applied (to letter selection), the correlational improbability goes to one. The log of one equals zero. This subsequently pushes the second part of the equation to zero and this in turn "zero's"(verb) the entire equation, giving us a zero information reading. The monkey of course, has no Shakespeare sonnets in his head that he is intending to type an informational correlate of. There is no rational reason to expect a monkey to produce Shakespeare either in his mind or on paper. The monkey is instead just typing at **random**.

It should be noted that due to "combinatorial explosion" Shakespeare patterns of any reasonable length (more than 68 letters)[1] *the probability (of its chance occurrence) quickly becomes unreasonably low such that the frequency of its occurrence/recurrence extends far beyond the lifetime of our universe. Such long sequences trigger a design inference (cheating?) and rule out the "production by chance"/monkey hypothesis -- assuming that "monkey" here is a metaphor for "chance."*

Example Three: Random Mutations.

Here again we have the question of randomness and its relation to information production. This question therefore applies not only to *random* mutations but also *random* gene duplications, *random* genetic drift, *random* genetic frame shifts, *random* Hox-gene shuffling and *random* neo-darwinian biological variations/changes of any kind. The problem is that unguided random events are always uncoordinated and therefore uncorrelated. Without organized correlations (fig #1) there is no "Brookfield information." There is just useless gibberish - in this case biological gibberish. DNA is said to contain information because DNA coding *correlates* with biological form and function. The source of this co-ordination cannot possibly be randomness or filtered (NS) randomness. Darwinian mechanisms therefore provide no answer as to the origin of biological information.

Example #4: DNA -- As noted above DNA coding is both improbable and correlated – both complex and specified. DNA subsequently contains "real information" or "Brookfield information." the only known source of such information is intelligent design.

Conclusion:

A house is a three dimensional structure. One cannot reduce a house to a two or one-dimensional structure without performing a destructive over-reduction. Any such "house" is definitely "not a home." I am proposing that "information" similarly cannot be reduced to a one-dimensional string of characters and that two-dimensions are required to describe "information." The two components of my proposed equation (above) correspond to the two-dimensions of information necessary for a complete, but nonetheless parsimonious, mathematical formulation of "information." Moreover, this formulation integrates Shannon's "information" and Dembski's "complex specified information" into a single formula.

1. See William A. Dembski and Robert J. Marks II, "Conservation of Information in Search: Measuring the Cost of Success," IEEE Transactions on Systems, Man and Cybernetics A, Systems & Humans, vol.5, #5, September 2009, pp.1051-1061 "A. Monkey at a Typewriter" Pg. 1058