

“ENTROPY, FREE ENERGY AND INFORMATION IN LIVING SYSTEMS”

Andy C. McIntosh

School of Process, Environmental and Materials Engineering

University of Leeds, LEEDS, U.K. LS2 9JT

a.c.mcintosh@leeds.ac.uk

Abstract

In this paper we consider the concept of information and consider the various alternatives for its definition. The reductionist approach has been to regard information as arising out of matter and energy. Coded information systems such as DNA are regarded as accidental in terms of the origin of life and that these then led to the evolution of all life forms as a process of increasing complexity by natural selection operating on mutations on these first forms of life. Thermodynamicists have long been aware that organisational systems are inherently systems with low local entropy and have argued that the only way to have consistency with an evolutionary model of the Universe and common descent of all life forms is to posit a flow of low entropy into the earth's environment and in such a model they suggest that islands of low entropy form organisational structures found in living systems.

There is a third alternative explored in this paper which proposes that information is in fact non-material and that the coding of DNA and all living systems is not defined at all by the biochemistry or physics of the molecules used to store the data. Rather than matter and energy defining the information sitting on the polymers of life, the reverse is the case. Information has its definition outside the matter and energy on which it sits and furthermore constrain it to operate in a highly non-equilibrium environment thermodynamically. This approach resolves the thermodynamic issues and invokes the correct metaphysical paradigm for understanding the vital area of thermodynamic / organisational interaction, which despite the efforts from alternative paradigms has not given a satisfactory explanation of the way information in systems operates.

1. Introduction

The concept of information has been a major issue since the discovery by Francis Crick and James Watson of the coding structure of DNA in 1953. Crick himself stated¹

If the code does indeed have some logical foundation then it is legitimate to consider all the evidence, both good and bad, in any attempt to deduce it.

This was stated in the context of the discovery that triplets of nucleotides running along the rungs of the double helix molecule of DNA carry information to code for a specific amino acid which then makes up the proteins of the living organism. Crick was always of a reductionist mind set and had no sympathy with any approach which regarded the coding as essentially an expression of a non material intelligence transcendent to the polymer itself, and the above statement in its original context is most definitely not advocating an exploration of information in any other paradigm than a purely materialist approach. However it is significant because it shows that scientific investigation can be trapped by only considering one pathway – what if the search for a ‘logical foundation’ advocated by Crick, actually leads one to the *edge* of the material region of scientific enquiry?

Stephen J Gould wrote of non overlapping magisteria², often referred to with the acronym NOMA, in order to resolve the issues of how to approach science describing the physical realm and the metaphysical / philosophical concepts describing the realities dealing with the non material. This is diagrammatically shown in figure 1.

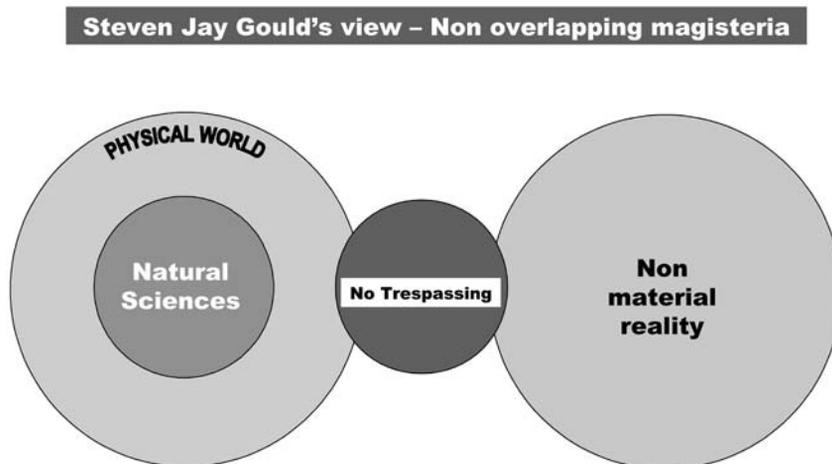


Figure 1. Stephen Jay Gould's non-overlapping magisteria (NOMA) view of reality.

In this paper we argue that such an approach to reality means that when one comes to investigate the area of information and software, and also mind and consciousness, this incorrectly locks the investigator into a materialism which actually denies

the most persuasive explanation of the intricate systems which have come to be understood in recent years. The antithesis to Gould's approach is illustrated in figure 2. It is argued that there is a legitimate realm where information, mind and consciousness lie – this area is undoubtedly interacting with the physical realm but is not entirely controlled by it. We are here not talking about spiritual matters, simply the area where thoughts, logic and mind exists and where the importance of arrangement rather than matter itself is dominant – as in the sequencing of the nucleotides in DNA.

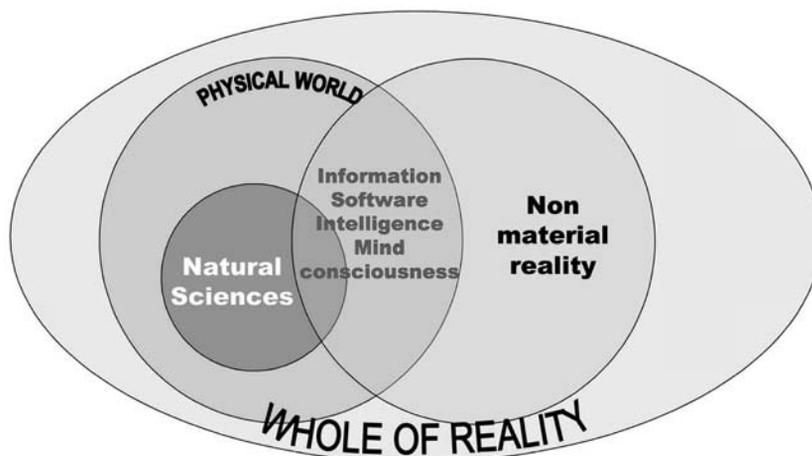


Figure 2. The view of reality advocated in this paper which defines information as constraining the matter and energy it sits on, but not defined by it.

The paradigm adopted here is the assumption that information is essentially defined as non material but profoundly influences the material in which it is found, in a similar way that software is essentially coded non material instructions but nevertheless controls the hardware of a computer. It should be emphasised that this is not a license for any lazy approach, whereby anything which cannot be understood is put metaphorically into a box labelled 'non material and not to be further investigated'. This is no 'god of the gaps' thesis. Indeed, once adopted, this approach opens out a whole raft of new research routes which properly explain the control of living systems. A far more profound methodology is in view. What is being advocated here is an entirely different paradigm whereby the non-material message / logos is accepted as being of an origin outside the area of physical investigation, but that its effect can readily be seen in the organisation of the molecular machinery in living organisms. Rather than the material and energy forming the information system as advocated by evolutionary philosophy, the non material informational message expressed in the coded ordering of nucleotides is actually the mechanism of constraining the material itself. It is the information which *organises* and *constrains*

the biopolymers. It is a known feature of living systems that they are information-rich and it is this that is more and more being recognised as the cause of their great efficiency³. Rather than the intricate machinery for such systems evolving from simpler systems, we show in this paper that the message/information itself is sitting on biochemical molecular bonds which are in a significantly raised free energy state. Understanding the thermodynamics of this machinery shows that it is thermodynamically impossible both to form such machinery without intelligence (abiogenesis) and that the laws of thermodynamics prohibit any formation of new machinery which is not there already or latently coded for in the DNA (evolutionary development). A hierarchical system of information is involved in living systems (see figure 3).

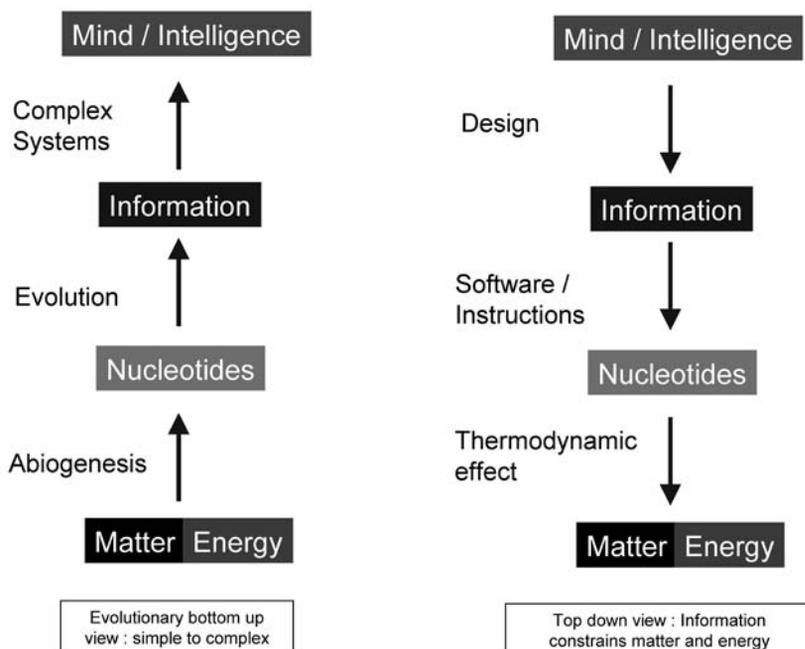


Figure 3. Hierarchical layering of living systems. The evolutionary view is that abiogenesis led to the forming of the DNA code which then led to the emergence of complex information systems and intelligence. However the top down view regards information in a similar way to the software instructions of a computer. The instructions organise the nucleotides and control the biopolymers to be in a highly non equilibrium state.

Furthermore recent research has confirmed that non-coding parts of DNA previously thought to be 'junk DNA' are in fact not to be regarded as such⁴. More research is now coming to light⁵ that the very folding of proteins carries with it a separate form of information transfer. This intertwining of information and matter

lies at the heart of what is life itself, and fundamentally changes our view of how to understand living systems.

2. Entropy and thermodynamics applied to DNA

In an earlier paper⁶ we have discussed the problem faced by Dawkins and others of arguing that the emergence of DNA (see figure 4) could come about by a random gathering together of the sugar phosphates and nucleotides. We briefly summarise the arguments here.

Dawkins⁷ argues that

Information is what enables the narrowing down from prior uncertainty (the initial range of possibilities) to later certainty (the 'successful' choice among prior probabilities) natural selection is by definition a process whereby information is fed into the gene pool of the next generation.



Figure 4. Sample DNA genetic code with main and complementary strands

In Dawkins' scenario he has to hunt for new information to counter the 'narrowing effect of natural selection'. His definition of 'information' is simply that natural selection is the favourable advantage of random mutations in one generation making that alteration more likely to survive and be thus more prolific in the next generation. The narrowing effect is that the number of options to choose from is reduced since once the selection is made, the original gene pool is reduced. The answer suggested here is to posit the topping up of the gene pool by the very mutations themselves.

There are two major obstacles to such a proposal. First the code is highly sequence specific. Each triplet of nucleotides codes for a specific amino acid and the protein formed from these requires a specific sequence of such amino acids. For example there are enzymes which are *specifically* assigned to nucleotide excision repair – they recognise wrongly paired bases in the DNA nucleotides (Adenine (A), Thymine (T), Cytosine (C) and Guanine (G)) connecting the two deoxyribose sugar-phosphate strands (see figure 4). This means that mutations are generally corrected (Jackson⁸ and de Laat et al⁹), so that even if speciation does occur due to slight modifications and adaptations of the phylogeny, any serious departures in the genetic information would be acted against by the DNA's own repair factory. Mutations do not increase

information content – rather the reverse is true. The flightless Galapagos Cormorant is a classic example. Evidently repair by the above techniques was not possible, and the genetic defect has persisted, such that information has certainly been lost, and the gene pool (in that case irrevocably) reduced. At the very least Dawkins' assertion at the end is misleading, for it suggests there is a natural *source* of new information which experimental observation denies. Natural selection cannot be redefined and is not the handmaid of macro evolution.

The second obstacle is a more fundamental issue. At the molecular level, the laws of thermodynamics themselves do not permit step changes in the biochemical machinery set up for a particular function performed by the cells of living organisms. That is there is an 'uphill' gradient in the formation of any of the molecular bonds in the nucleotides and most of the proteins. There is no chemical drive to form these, rather there is a move to equilibrium away from their formation.

2.1. The second law of thermodynamics

A number of statements of the second law are equivalent. One form of the second law is "The amount of energy available for useful work in a given isolated system is decreasing. The entropy (dissipated useful energy per degree Kelvin) is *always* increasing."

Thus according to the second law, heat always flows from hot to cold. In the process it can be made to do work but always some energy will be lost to the environment, and that energy cannot be retrieved. Water flows downhill and loses potential energy which is changed into kinetic energy. This can again be made to do work (as in a hydroelectric power plant). However some energy will be lost such that if one was to use all the energy generated to pump the same water back up to its source, it would not reach the same level. The difference of original potential energy to that corresponding to the new level, divided by the temperature (which in that case is virtually constant) is the entropy of the system. Such a measure will always give an entropy *gain*.

There is no known system where this law does not apply. The fact that the entropy of a given isolated system increases, effectively brings with it an inevitable decline in usefulness of all systems.

Non isolated systems

When one considers non-isolated systems where heat transfer can take place at the boundary, some have argued that by adding energy in to the original system then one should be able to reverse the trend of entropy increase. However as ref 6 shows, this is not the case. Adding energy without an existing machine to make use of that increase, simply leads to the heating up of the surroundings faster than

would otherwise have been the case. Only by having in existence a machine for capturing the incoming energy can further useful work be achieved.

Prigogine¹⁰ and others (see for instance Styer¹¹) have proposed that there is meaning in the non material arrangement and organisation of systems and refer to an organisational entropy or 'logical' entropy. They propose the addition of other entropies which could then feed negative entropy into a given (non isolated) system. Consequently the total entropy is considered to be

$$ds = ds_T + ds_{\text{logical}} \quad (1)$$

Where ds_{logical} represents the ordering principle or 'logical' negative entropy which gradually seeps in to the system. Thus even though ds overall is increasing with the thermal entropy ds_T positive, the presence of ds_{logical} coming in at the boundary ensures locally the low entropy needed to spark evolutionary development. Styer¹¹ speaks of a net entropy flux at the earth which would then be the source of evolution of early prokaryotes to eukaryotic individuals.

Thus complexity and the ordering principle is predicated on the notion that information can gradually increase from a random state. Again this is flawed for two reasons. First as stated earlier, no flux of energy from outside the system can be made to do work within the system unless there is the necessary machinery to capture this energy. Secondly the information required is not defined in purely thermodynamic terms or even in any ordered code such as the DNA. Gitt¹² has shown that information is hierarchical in at least five levels. Two important levels are code (or language) and message which uses the coded communication system. Neither of these can actually be thought of as arising simply from a flux of entropy locally. However the information does sit in living systems where there are thermodynamic effects.

3. Free energy and Machines

Free energy

The Gibbs free energy g is defined as the net energy available to do work. It effectively takes away the unusable lost energy (associated with entropy) from the enthalpy h (which can be regarded as the total thermodynamic energy available). Thus

$$g = h - Ts, \quad \text{and} \quad \Delta g = \Delta h - T\Delta s \quad (2a,b)$$

Machines

As a consequence of the second law of thermodynamics applied to non-isolated systems⁶ one can state that the following applies concerning the spontaneity of chemical reactions:

$$\begin{aligned}
 \Delta g < 0 & \text{ favoured reaction – Spontaneous} \\
 \Delta g = 0 & \text{ reversible reaction – Equilibrium} \\
 \Delta g > 0 & \text{ disfavoured reaction – Non-spontaneous}
 \end{aligned}
 \tag{3}$$

A positive free energy device cannot arise spontaneously. It always requires another operational machine to enable the free energy to be loaded / ‘primed’, ready to do work. This is best illustrated in the example below (figure 5) of a compressed spring, and we now define a *machine* as a device which can locally raise the free energy to do useful work.

Even if material exchange was involved (and one had a completely open system), no amount of matter or energy exchange without information exchange would alter the fundamental finding (eqn (3)) concerning spontaneity of chemical reactions.

The use of these type of free energy arguments now applies to non isolated systems (that is where heat is allowed to cross the boundary of the system) as well as to a system completely isolated with no contact with the surroundings. One can now consider what would happen if energy were directed to a system (as in section 2) and it is evident that without a machine the free energy is not increased, *so that no new machine will arise simply by adding random energy into an existing system.*

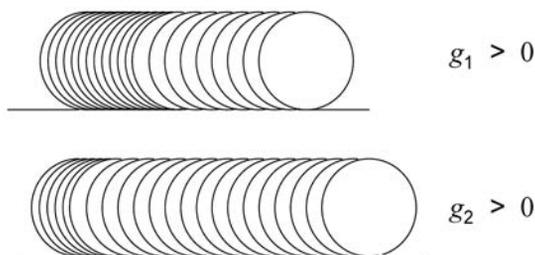


Figure 5. All natural molecule formations are like coiled springs such that if one lets the system ‘go’ they ‘relax’: $\Delta g < 0$ (due to $g \equiv h - Ts > 0$). To ‘set’ such springs requires an *initial input* of ordered energy by another machine. In this example the internal energy in the 2nd case g_2 is less than the first case g_1 .

Consequently to suggest that reactions on their own can be moved against the free energy principle is not true, since they could not be sustained. The DNA molecule along with all the nucleotides and other polymers could not change radically such that a low entropy situation would emerge. To alter the DNA constituents from one stable state say to another representative state with a distinct improvement

cannot be done by natural means alone without additional *information*. The laws of thermodynamics are against such a procedure.

Entropy and Open Thermodynamic Systems



Photosynthesis in a living plant



The energy from the sun is absorbed (along with carbon dioxide and water) by plants through photosynthesis. The chlorophyll of the leaf acts as a catalyst to the biochemical reaction:



Carbon Dioxide Water Sunlight Sugar Oxygen
 Energy + **Information** → Locally reduced entropy (Increase of order)
 (or teleonomy)

Energy on its own **does not** lead to decrease in Entropy

Figure 6. Photosynthesis in a living plant requires energy input, but the energy flux on its own would do nothing unless there was a machine already present (a free energy device) to enable the system to do work using the sunlight.

Put another way the carrier molecules of the information in living systems are kept in a disequilibrium state by the very presence of information. They would fall apart to a disordered equilibrium state were it not for the information in the system making them act in this way.

4. A new approach: Entropy constrained by functional information

We propose a different treatment which quantifies the effect of functional information in a system. This approach recognises Gitt's important deductions concerning real information systems being impossible to define in terms of matter and energy alone¹². However one can recognise the effect of machines / information systems (that is teleonomy) being present in exactly the same way as a digitally controlled machine (i.e. a computer) is operated by software. The high level program controls a set of electronic switches on a micro chip which are set in a certain predefined pattern (see right hand diagram in figure 3). Thus the logical entropy (the switching of the micro chip in the analogy) rather than being the *source* of the information should be thought of as the *effect* of information carrying systems.

Only with the presence of a free energy device (a machine) already existing will an energy flux outside the system do useful work and enable a local lowering of the entropy of the system. This is illustrated for photosynthesis in figure 6.

In this approach it is expected that there will be levels of information, and in particular language (code) and semantics (meaning). This approach^{6,13,14} leads to the proposition that machinery and information are closely intertwined in living systems.

For a pure materialist there may be a natural reticence to adopting such an approach because of his presuppositions, but the evidence of the thermodynamics of living systems supports the view that it is *information* in living systems that controls the thermodynamics, and not the other way round.

5. Conclusions

Three views of informational reality (ontology) are considered in this paper. The first is that matter and energy is all there is. This is the materialistic view of information (Dawkins (Oxford), Jones (University College, London), Atkins (Oxford) and others). Such authors argue that functional non-material information and design are an illusion. In reality (in their view) matter and energy is all that there is in the Universe. Patterns only have meaning in a reductionist sense and do not carry any non material 'value'. The second scenario is a variation of the bottom up approach. Information is regarded as non-material but has arisen out of matter and energy. This is the view of Prigogine¹⁰, Yockey¹⁵ and Wicken¹⁶ and many other authors. Both these approaches are flawed on two counts. Firstly they ignore the fact that real information systems are not defined by the codes and languages they use and that the arrangements of the physical objects used in the system (for DNA, the nucleotide triplets) is in a specified order. So even a non-materialist, but bottom up approach does not have the means to derive specificity¹⁷ in the ordering arrangement of the nucleotides in DNA. Secondly a more subtle point but a very important one is that there is an impossible thermodynamic barrier to such an approach. The information system in living systems is mounted on molecules with a raised free energy such that the carriers of information would fall apart into equilibrium chemistry were it not for the information present. It is this barrier which shows that a top down approach is the only way to understand information in living systems.

The third view then that we have proposed is the top down approach. In this paradigm, the information is non-material and *constrains* the local thermodynamics to be in a non-equilibrium state of raised free energy. It is the information which is the active ingredient and the matter and energy are passive to the laws of thermodynamics within the system.

References

- [1] Crick, F. *On the Genetic Code*, Nobel Lecture, December 11, 1962 and also recorded in, *Science* 139(3554), 8 February 1963.
- [2] Gould, S.J. *Non Over-lapping Magisteria*, *Skeptical Inquirer*, July August 1999, 55-61.
- [3] Vincent, J.F.V., Bogatyreva, O.A., Bogatyrev, N.R., Bowyer, A. and Pahl, A.K. *Biomimetics: its practice and theory*, *Journal of the Royal Society*. Interface 3, 471-482, 2006. In particular see discussion on p. 477 of nature's vast resources of information and small energy useage by comparison with man made technology.
- [4] Birney, E. et. al., "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project", *Nature*, 447, pp. 799-816, 14th June 2007.
- [5] Tramontano, A., *The ten most wanted solutions in protein bioinformatics*, Chapman and Hall, 2005. See chapter 4 "Problem 4 Protein Structure Prediction, pp. 69-88.
- [6] McIntosh, A.C. "Functional Information and Entropy in living systems", pp. 115-126, *Design and Nature III: Comparing Design in Nature with Science and Engineering*, Vol. 87 of *WIT Transactions on Ecology and the Environment*, Editor Brebbia, C.A., WIT Press, 2006.
- [7] Dawkins, R. *The Information Challenge*, pp. 107-122 (quote is from p. 121), Chapter 2.3 of *A Devil's Chaplain; Selected Essays by Richard Dawkins*, Ed. Latha Menon, Phoenix, 2003.
- [8] Jackson, S.P., "Sensing and repairing DNA double strand breaks", *Carcinogenesis*, Vol. 23, No. 5, 687-696, OUP, May 2002.
- [9] de Laat W. L., Jaspers, N.G.J., and Hoeijmakers, J.H.J., "Molecular mechanism of nucleotide excision repair", *Genes and Development*, Vol. 13, No. 7, pp. 768-785, April, 1999.
- [10] Nicolis, G. and I. Prigogine. *Exploring Complexity: An Introduction*, W.H. Freeman, New York, 1989.
- [11] Styer, D.F. "Entropy and Evolution", *American Journal of Physics* 76 (11), 1031-1033, 2008.
- [12] Gitt, W, "Information: the third fundamental quantity", *Siemens Review*, Vol 56, Part 6, pp.36-41, 1989.
- [13] Trevors, J.T. and Abel, D.L., "Chance and necessity do not explain the origin of life, *Cell Biology International* 28, 729-739, 2004.
- [14] Abel, D.L., "The Capabilities of Chaos and Complexity", *Int. J. Mol. Sci.*, 10, 247-291, 2009.
- [15] Yockey, H., *Information Theory and Molecular Biology*, Cambridge University Press, 1992.
- [16] Wicken, J.S., *Evolution, Thermodynamics and Information: Extending the Darwinian Program*, Oxford University Press, May 1987
- [17] Meyer, S.C., "The origin of biological information and the higher taxonomic categories", *Proceedings of the Biological Society of Washington*, 117(2), 213-239, 2004.