

# Living *is* information processing: from molecules to global systems.

Keith D. Farnsworth<sup>1</sup>, John Nelson<sup>1</sup> & Carlos Gershenson<sup>2</sup>

<sup>1</sup> School of Biological Sciences, Queen's University Belfast, UK, BT9 7BL  
{k.farnsworth,John.Nelson}@qub.ac.uk

<sup>2</sup> Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas,  
Universidad Nacional Autónoma de México.  
A.P. 20-726, 01000, México, D.F., México  
cgg@unam.mx

**Abstract** We extend the concept that life is an informational phenomenon, at every level of organisation, from molecules to the global ecological system. According to this thesis: (a) living is information processing, in which memory is maintained by both molecular states and ecological states as well as the more obvious nucleic acid coding; (b) this information processing has one overall function - to perpetuate itself; and (c) the processing method is filtration (cognition) of, and synthesis of, information at lower levels to appear at higher levels in complex systems (emergence). We show how information patterns, are united by the creation of mutual context, generating persistent consequences, to result in 'functional information'. This constructive process forms arbitrarily large complexes of information, the combined effects of which include the functions of life. Molecules and simple organisms have already been measured in terms of functional information content; we show how quantification may be extended to each level of organisation up to the ecological. In terms of a computer analogy, life is both the data and the program and its biochemical structure is the way the information is embodied. This idea supports the seamless integration of life at all scales with the physical universe. The innovation reported here is essentially to integrate these ideas, basing information on the 'general definition' of information, rather than simply the statistics of information, thereby explaining how functional information operates throughout life.

*Keywords:* complex system; entropy; biocomplexity; evolution; network.

## 26 **1 Introduction: what is life?**

27 The question ‘what is life’ is one of the oldest in philosophy, deeply mysterious and still fascinating. Not  
28 only is it fundamental to biology, it has challenged and extended physics, metaphysics, the human sciences  
29 of medicine and psychology, the arts and even spiritual thinking. But efforts to answer the question have  
30 generally been constrained by disciplinary boundaries or within an organizational scale of life, leading to  
31 several apparently separate answers. The aim of this paper is to unite these by considering life as a whole,  
32 simultaneously at every organizational level (from molecule to global ecosystem). This integration uses the  
33 concept of life as information processing for a unifying principle.

34 During the second half of the 20th century, the paradigm that ‘life is chemistry’ ([Kornberg, 1991](#)) was  
35 especially influential in understanding living processes at the sub-cellular level. As increasingly complicated  
36 networks of molecular interactions were recognised, the need for a formal understanding of their organi-  
37 zational structures developed into systems biology, which now extends beyond the cell ([Kohl et al., 2010](#)).  
38 At the same time, but largely unrelated, theoretical ecology developed into a form of cybernetics: the study  
39 of self-regulating systems, moving chemical substances through networks of populations and communities.  
40 The complex networks of the cell’s biochemistry were paralleled by complex webs of interactions among  
41 organisms: the elaborate complexities of the ‘-omics’ were matched by those of biodiversity as we realised  
42 that the estimated 15 million species (8.7 million eukaryotic ([Mora et al., 2011](#)) plus 6 million prokaryotic  
43 ([Curtis et al., 2002](#))) are all connected to one-another in networks of community interactions. Observing  
44 that these complex networks may be two manifestations of a common feature of life, we now propose a  
45 unifying model in which interactions among molecules, cells, organisms and populations all amount to in-  
46 formation processing through a hierarchy of functional networks - molecules in cells, cells in organisms and  
47 organisms in communities, which compose the biosphere. This model, which extends recent developments  
48 in systems biology ([Maus et al., 2011](#)) is intended to integrate through all life over its entire history.

49 Biologists know that information is crucial to life, pointing to its role in DNA for maintaining the design  
50 of organisms over repeated generations and an understanding of information in protein structure has a long  
51 history (see e.g. [Yockey et al., 1958](#)). A cybernetic view goes further to claim that information processing,  
52 carried out in the medium of biological chemistry, is what life actually *is*. By information processing we  
53 mean any logical combination of information having the result of producing information and we shorten  
54 this to ‘computation’. The idea that ‘living is computing’, pioneered by theorists such as [Galtin \(1972\)](#) has  
55 been popularised by [Bray \(1995, 2009\)](#), but so far, it has been contained within cellular biochemistry (with  
56 computation by neural networks the obvious exception). Our aim is to show how well the whole of life can

57 be viewed in this way as an integrated information processing system: all cells working together. This view  
58 seamlessly connects with the concept of information as one of three elemental components of existence  
59 (with space/time and matter/energy) which has grown within physics over the past several decades, accom-  
60 panied by a new philosophical position which places information at the core of determining reality (termed  
61 ‘Informational Structural Realism’ by Floridi (2003)). Every aspect of life may be regarded as a product and  
62 elaboration of the physical world, clearly made of the same matter and energy, ordered in space and time as  
63 is every physical system. What makes life special is not the material brought together to take part in living,  
64 it is the functional information that orders matter into physical structures and directs intricate processes  
65 into self-maintaining and reproducing complexes. In the information model of life, this definitive process  
66 (termed autopoiesis by Maturana and Varela (1980) consists of a system of structural elements continually  
67 replacing themselves to maintain the living system by following a program of instructions that both makes  
68 their information-rich structure and is instantiated within it. Significantly, this fundamental feature of life is  
69 true at every organizational scale, not only at the cellular level.

## 70 **2 Information Concepts**

71 According to the ‘diaphoric definition of data’ (Floridi, 2003, 2005), a binary bit (the unit of information)  
72 is a single difference. For example, a digital monochrome image of  $k$ -pixels instantiates no more than  
73  $k - 1$  differences. When the image carries a meaningful picture, it instantiates fewer than the maximum  
74 number of differences, so can be compressed by recording only the differences where black changes to  
75 white. The maximally compressed image instantiates  $k - n$  bits ( $n \geq 1$ ) and this is termed the Algorithmic  
76 Information Content (AIC) (Chaitin, 1990). The same applies not just to representations, such as images,  
77 but to real physical objects: a compressible pattern of differences makes an object what it is. This refers  
78 not to a description, but to the physical object itself, giving a definition of physical information as a pattern  
79 of difference: the algorithmic information embodied by an object so as to give it form. Information in this  
80 sense, selects the elementary particles of the object and specifies the locations of these in space and time  
81 (under quantum-theoretic constraints). The minimum description of the object is the AIC embodied in both  
82 this physical configuration of particles and the nature of each (Pauli’s exclusion principle ensures these  
83 are different). On a technical note, AIC is known not to be strictly computable (Li and Vitányi, 2008), but  
84 an effective substitute is available in the Computable Information Content for empirical studies needing to  
85 compute it (see e.g. Menconi, 2005).

86 For most practical purposes, in describing an object, we would consider higher levels of abstraction,  
87 such as a pattern of atoms, molecules, cells, tissues, or components, etc.. Again, for most practical purposes,  
88 we are concerned not with the total AIC instantiated in an object, but with the *functional information content*  
89 (FIC), which is the part of AIC which can cause a persistent change of information in any part of the system.  
90 As an illustration, two seemingly identical metal keys will be different in detail (at the small scale), but may  
91 both function to open the same lock: their functional information defines their shape as fitting the lock.  
92 This is obviously pertinent to biology through the lock and key analogy of messenger molecules, but also  
93 describes functional equivalence among all kinds of biological molecules; among cells of the same type and  
94 state in the body; and among organisms of the same function in an ecosystem. FIC can be quantified, as  
95 demonstrated at the nucleotide level by [Jiang and Xu \(2010\)](#), who calculated it as the minimal amount of  
96 genomic information needed to construct a particular organism. We hope to apply this idea to structures of  
97 biological information, other than the genetic.

98 In the field of Biosemiotics, pieces of functional information are regarded as symbols (see [Favareau,](#)  
99 [2009](#)), but we wish to focus on the functioning of information, rather than its communication. For this,  
100 we take the idea of *function* from [Szostak \(2003\)](#), seeing it as what makes systems, including biological  
101 ones, operate, in the sense of an operational explanation of function ([Neander, 2011](#)). The definition of  
102 ‘function’ has been debated among philosophers for several decades and deserves some attention here.  
103 [Cummins \(1975\)](#) proposed that function is an objective account of the contribution of a system component  
104 to the ‘capacity’ of the system. Crucially, for Cummins, the capacity (meaning capability) of a system  
105 is explained in terms of the capacities of the components it contains, and how they are organised. This  
106 concept explicitly matches the understanding that functional information is to be found in the component  
107 parts and the way they are organised into a whole. But it has been criticised, especially for its permitting  
108 what appear to be unintended consequences as functions (a frequently cited example being that dirt in a  
109 pipe may ‘function’ as a valve ([Griffiths, 1993](#))). One of the solutions to this, at least for organisms, is to  
110 recognise that natural selection tends to eliminate potential functions of components if they do not contribute  
111 to the biological fitness of the system of which they are a part. This qualification was taken up by [Neander](#)  
112 ([1991](#)), by developing a biologically-based etiological theory. Whilst appealing, this cannot be used for all  
113 biological systems, such as ecological communities, for which evolution by natural selection has not been  
114 established, so to be general, we are forced back to the systemic theories of function. However, Darwin’s  
115 theory is a special case of a more general principle of selection in which the attribute of persistence is the  
116 superset of biological fitness (e.g. [Kauffman, 1993](#)). Thus we tentatively offer a definition of function that

117 is systemic and in the spirit of established etiological definitions, but not reliant on Darwin's theory. It is  
118 that any attribute A of a component C of a system S that causes an effect E such that S persists longer or in  
119 a wider range of conditions than without it, is a functional attribute of C. Then the functional information  
120 instantiated by C is that which establishes A, leading to the persistence in form of S, hence the persistence  
121 of information instantiated by S.

122 Szostak's (2003) mathematically amenable definition allows for a quantification of the effectiveness  
123 with which information enables a system to perform non-random actions; at least one of which will be  
124 self-replication. From here on, we shall use the general term '*effective information*' for that which causes  
125 a persistent change, so has an effect in the wider system and reserve the term '*functional information*' for  
126 effective information which plays a role in supporting life. We note that at the specifically nucleotide level,  
127 since evolution selects for function, non-functional information will be lost from biological systems over  
128 evolutionary time (this was demonstrated by Schneider's (2000) 'evolutionary program'). However, non-  
129 functional information is continually introduced by random processes, especially at higher (e.g. ecological)  
130 levels, so non-functional 'noise' may be expected and should be discounted in the quantification of FIC.

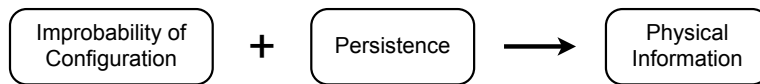
131 We take as axiomatic that information is instantiated in matter through the particular arrangement of  
132 its components in space and time. This arrangement defines a unique relationship among the components,  
133 which can only instantiate information if it is stable and therefore persists as a configuration in space over  
134 a line in time. When two or more such configurations are brought into association, there is a combined  
135 arrangement, which if persistent, also instantiates information: that of both components *plus* that of their  
136 association. The Shannon information (Shannon, 1948) of the combined configuration is given by the prod-  
137 uct of probabilities of each component configuration (less any mutual information). Thus the 'surprise' in  
138 finding this new whole is in general greater than that for each of its component parts. Nested construc-  
139 tion of increasingly complicated configurations of matter may proceed this way and thereby constitute an  
140 increase in information content in the Shannon sense (Shannon, 1948). Most significantly, when configura-  
141 tions combine into stable forms, they do so by presenting context for one another: the information of each  
142 is functional information for the other, enabling greater function than that of the sum of parts.

143 The functional meaning of information was defined conceptually by MacKay (1969) who referred to  
144 information as "a distinction that makes a difference" and later Bateson (1972) more famously called in-  
145 formation "a difference that makes a difference", this idea was then taken up by Hopfield (1994). In this  
146 interpretation, information is defined through its interaction with something (including other information)  
147 to create a non-random effect, hence it is context dependent. Bates (2005), quoting earlier works, defines

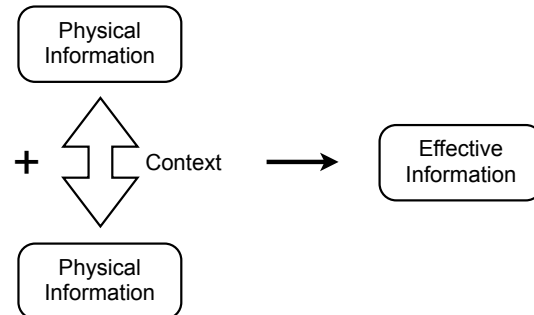
148 information as: "the pattern of organization of matter and energy". This definition peculiarly addresses ef-  
149 fective information. Patterns of organization are the alternative to randomness: patterns show either order  
150 (characterised by symmetry) or complexity (broken symmetry). [Schrödinger \(1944\)](#) realised that symmet-  
151 rical order was insufficient to account for the genetic information coding life, concluding that it must be in  
152 some aperiodic (non-symmetrical) molecule (well before the discovery of DNA). The required organized  
153 aperiodicity is commonly known as 'complexity'; a defining characteristic of which is a high capacity for  
154 effective information. [Adami et al. \(2000\)](#) subsequently showed how all biological systems are complex  
155 systems in this scientific sense.

156 These concepts are brought together in Figure 1 which shows three levels of information concept in the  
157 formation of life. On level 1, physical information is understood as the result of an improbable (following  
158 Shannon's insight) and persistent configuration of energy and/or matter in space and time. In level 2, ef-  
159 fective information is defined through consequence: a contextual relation is made among at least two such  
160 configurations (now considered as information and termed 'infons'). This synthesis through mutual context  
161 is exemplified by a lock and key enzyme interaction. Level 3 takes this further to capture the idea that a  
162 large number of contextual interactions structure an assembly of infons into a complex system; exempli-  
163 fied by a molecular network inside a cell. Not shown is the hierarchical concept that such systems can be  
164 the component parts of super-systems, enabling an unbounded construction of nested complexity, in which  
165 information at higher levels, but not present at lower levels, can be defined and measured as emergent ([Ger-  
166 shenson and Fernández, 2012](#)). That is the way life appears under observation, exemplified by the notional  
167 hierarchy in figure 2 and table 1.

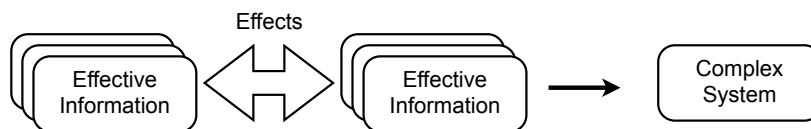
## Level 1



## Level 2



## Level 3



**Fig. 1** Three levels of information concept explained in the text: at level 1, information is a pattern of difference; at level 2, information becomes effective through context and at level 3, 'packages' of effective information combine, affecting one another to form a complex system that computes.

168 2.1 Order from disorder: self-assembling structures

169 According to statistical mechanics, the organization of a system is the result of filtering, i.e. selecting a  
 170 particular configuration of system component states from all possible configurations and this filtering is  
 171 equivalent to investing the system with information, in the Shannon sense (Shannon, 1948) of reducing the  
 172 probability of its configuration. When the resulting organization causes sustainable self-assembly, using  
 173 active filtration from the wider environment, the system may be said to live.

174 It is most parsimonious to assume that the components of matter needed to constitute living organisms  
 175 were originally distributed in perfect randomness (disorder). Apparently, life alone creates life, but before  
 176 it appeared for the first time, individually persistent (non-transitory) stages of ordering among collections  
 177 of molecular components must have occurred. It is broadly understood that this develops through the spon-  
 178 taneous emergence of 'order out of chaos' (von Foerster, 1960; Prigogine and Stengers, 1984; Kauffman,

179 [1993](#))—in which chaos then referred to disordered randomness. This natural evolutionary phenomenon,  
180 which obeys the second law of thermodynamics, is very general. It amounts to the selection of more stable  
181 configurations from a set of random configurations, simply by virtue of their stability conferring greater per-  
182 sistence. Darwin's evolution by natural selection is a particular instance of this process, which also applies  
183 to resonance phenomena and crystal formation.

184 Life orders matter, but differs from a crystal in the following critical respects: (a) life is a dynamic  
185 pattern not a static one; (b) it is not regular, but rather is complex, meaning that it cannot be summarized in  
186 a short piece of information and (c) it manipulates its environment so as to make its persistence more likely.  
187 The vortex (e.g. a whirlpool) is an often cited example of a non-living system which displays some of these  
188 properties. It maintains itself as a dynamic pattern of matter, even though its constituent parts are constantly  
189 changing: molecules which pass through in a moment are replaced by others, but the pattern and therefore  
190 the structure-forming information is maintained. This is an example of a 'dissipative structure' defined  
191 and recognized as self-organizing by Prigogine ([1977](#)). By continually exchanging matter and energy with  
192 their environment, these dynamic structures are able to continually 'dissipate' entropy, with the effect of  
193 concentrating information. This information is instantiated in the form of the structure. Crucially the essence  
194 of these dissipative systems is organizational information, not substance, and the information they maintain  
195 has the special property of being that which is necessary for the self-maintenance.

196 Given the required material components and thermodynamic conditions, we see that information in  
197 the form of a pattern in matter can emerge spontaneously and maintain itself as long as these conditions  
198 allow. The next step is to ask if it can also create the components and maintain the conditions it needs to  
199 do this in a changing environment. If any pattern can achieve that feat, then it will be able to reproduce  
200 and ensure its persistence far longer than thermodynamics would otherwise allow. The ability of a system  
201 (any arrangement of matter) to remake itself is termed autopoiesis and this has been identified as one of  
202 the two necessary capabilities of anything living ([Maturana and Varela, 1980](#)). The other is cognition, more  
203 precisely, the detection and selection of particular elements from an environment of many random elements,  
204 which is a kind of information processing. [Bitbol and Luisi \(2004\)](#) showed that autopoiesis and cognition  
205 are separately necessary conditions for life, not inseparably linked as apparently first thought by Maturana  
206 and Varela ([1980](#)). They illustrated their point with reference to the autopoietic fatty acid cells, which [Zepik](#)  
207 [et al. \(2001\)](#) showed to achieve reproduction and self-maintenance by homeostatic processes autonomously  
208 generated from within. From this work, it became clear that for a system to live, it must have at least the  
209 following three properties: autopoiesis, cognition and an unbroken boundary to define its limits ([Bitbol and](#)



210 [Luisi, 2004](#)); this latter stops the ingredients of life from diffusing apart, rendering life's chemical reactions  
211 too rare to work as a whole. In practice, all known living systems are cellular<sup>1</sup> and indeed, the cell tegument  
212 has never been broken since the beginning of life—it has only been divided by repeated fission. Division  
213 among organisms is just an elaboration of division among cells. In this sense all life from its beginning,  
214 is unified as a set of cells, related through replication; all creating order from disorder, by cognition and  
215 autopoiesis.

216 The result of this long history of accumulating functional information in a population of diverging cell  
217 lines is illustrated in Figure 2 where the major developments are illustrated. By specializing into specific  
218 types, cells have found ways to more effectively live: colonies of specialist cells forming into the distinct  
219 tissues of separate organisms, organized into ecological communities, interacting, to the point of regulating  
220 the earth's geochemistry through a homeostatic network. All of this amounts to information processing—  
221 selecting molecules from the environment, ordering matter and controlling flows of matter and energy. The  
222 information needed to perform these functions is found distributed among the molecules within every cell:  
223 not just in nucleotides, but in all the proteins and messenger molecules, their interactions and locations in  
224 space. However, seeing life as a whole in space and time, from the first single cell to all extant life, implies  
225 an integrated system, for which hierarchical levels represent merely *observed* abstractions of organisational  
226 structure (see [Salthe, 1985](#)). Considering the whole living system from notional levels of biochemistry at the  
227 bottom to global ecosystem at the top, we may regard all but one of the levels in table 1 to be a model, the  
228 single exception being organisation into cells. Hierarchy theory recognises constraints imposed by higher  
229 levels on the lower, but also the constraint of possibilities from lower levels upwards. We understand the  
230 need for bounded cells as one of those possibility constraints and therefore see cells as the one exception -  
231 they are not merely a model level but one in the reality of life's organisation.

## 232 2.2 Biological systems as effective information

233 It is evident that the minimum functional information needed to constitute life is large (the smallest non-  
234 virus functional information content calculated so far is  $2.86 \cdot 10^6$  bits for *Holarctica* ([Jiang and Xu, 2010](#))).  
235 By current consensus, life emerged as an entropy-dissipating pattern which created and maintained a bound-  
236 ary through which trapped molecules were able to selectively interact with the wider environment ([Mo-  
237 rowitz, 1992](#); [Smith and Morowitz, 2004](#)). This cognitively filtering system also reproduced itself by growth  
238 and fission and all extant life followed via evolution ([Robertson and Joyce, 2010](#)). The resulting proto-cell

---

<sup>1</sup> Though some biologists may include viruses.

239 was a complex dynamic system in which information was held, not just in the component molecules, but  
240 also in the interactions among them. These interactions instantiated functional information because the  
241 molecules gave context to each other, thereby filtering out specifically functional interactions from the  
242 whole range of possibilities.

243 The cytoplasmic contents of cells are spatially structured so that the time and place of interaction is a  
244 necessary determinant of their effect. Because molecular components are distributed in a specific spatial  
245 pattern, their collective behavior is extended to form regions of coordinated, but different action over space.  
246 This instantiates functional information in spatial relations so that simple unitary systems (e.g. enzyme in-  
247 teractions) combine to exhibit complex behaviors which *appear* to be the product of more complicated com-  
248 ponents. The apparently spontaneous emergence of new information (Gershenson and Fernández, 2012), is  
249 in fact the revelation of that spatio-temporal information already present in the distribution of components  
250 and the network of signaling paths among them (a phenomenon first described by Turing, 1952). Any  
251 spatio-temporal information (coding the positions of system elements in time and space) that contributes to  
252 the emergent behaviors of the whole system, is effective information, and in life this is maintained by au-  
253 topoiesis. When a more complicated system is created from simple units in this way, it results in a new unit,  
254 the combination of these being the next tier in an hierarchy of complexity. It is by this nested hierarchical  
255 construction that the enormously complex machinery of life is brought into being.

256 Information is therefore not just stored in nucleotides: it is the whole biological system that embodies  
257 effective information, hence biocomplexity as a whole is the storage of effective information in living na-  
258 ture. Valentine (2003a) realised this and emphasised that biological complexity exists as a set of hierarchical  
259 levels, as we illustrate in table 1 (adapted from Farnsworth et al. (2012)). Spontaneous creation of effec-  
260 tive information from complex order is a signature property of such hierarchies: every level spontaneously  
261 *emerges* from the one below (Adami et al., 2000; Lorenz et al., 2011) - all the way up to global ecosystems.  
262 For this reason, even a complete description of genetic information fails to account for the full comple-  
263 ment of effective information in life, which is why seed-banks and zoos are no substitute for community  
264 conservation, as noted intuitively by Lee (2004) and Cowling et al. (2004). Indeed, ‘living information’ is  
265 only fully instantiated in dynamic, active systems capable of flexibly responding to environmental condi-  
266 tions. A common example is the gene-regulatory network, which apparently extracts maximum autopoietic  
267 complexity by functioning near criticality (Balleza et al., 2008), where information content is maximised  
268 (Gershenson and Fernández, 2012).

---

## 2.3 Quantifying Functional Information

Farnsworth et al. (2012) classified the total information content of any system into two distinct components:  $I_{tot} = I_F + I_R$ , where  $I_F$  is the functional information and  $I_R$  is the random information. Each of these terms can be quantified by the Algorithmic Information Content (Chaitin, 1990) if the term can be isolated.  $I_F$  could, in principle, be quantified by the ‘Effective Complexity’ (Gell-Mann and Lloyd, 1996, 2003), defined as the minimum description length of regularities, but only given prior knowledge about the regularities (see McAllister, 2003, for an expansion of this criticism). To describe life as information, we need a way to identify  $I_F$  without such prior knowledge, recognising that effect only results from the interaction of information and its context. In the special case of genomes, this is relatively trivial since almost all the information present is functional (Schneider, 2000). For quantification, Jiang and Xu (2010) defined ‘effective information’ as that part of the genome which is the minimum needed to reconstruct the organism. This meant estimating the functional (coding) fraction of the genome and (manually) compressing it to form the equivalent Algorithmic Information Content. In an application of Boltzmann’s entropy concept at the genetic level, Szostak (2003) defined ‘functional information’, in terms of a gene string, as  $-\log_2$  of the probability that a random sequence will “encode a molecule with greater than any given degree of function” - in other words a design brief, without implying a designer. In the case of genes, this ‘function’ may be thought of as the biochemical activity (for example a digestive enzyme’s catalytic rate) of whatever molecule is produced from reading the nucleotide sequence. This design-brief concept was developed to the ecosystem level of organisation by Farnsworth et al. (2012), who interpreted it as a set of ecological functions and related functioning to the information content of food-web networks.

## 3 The natural history of information processing

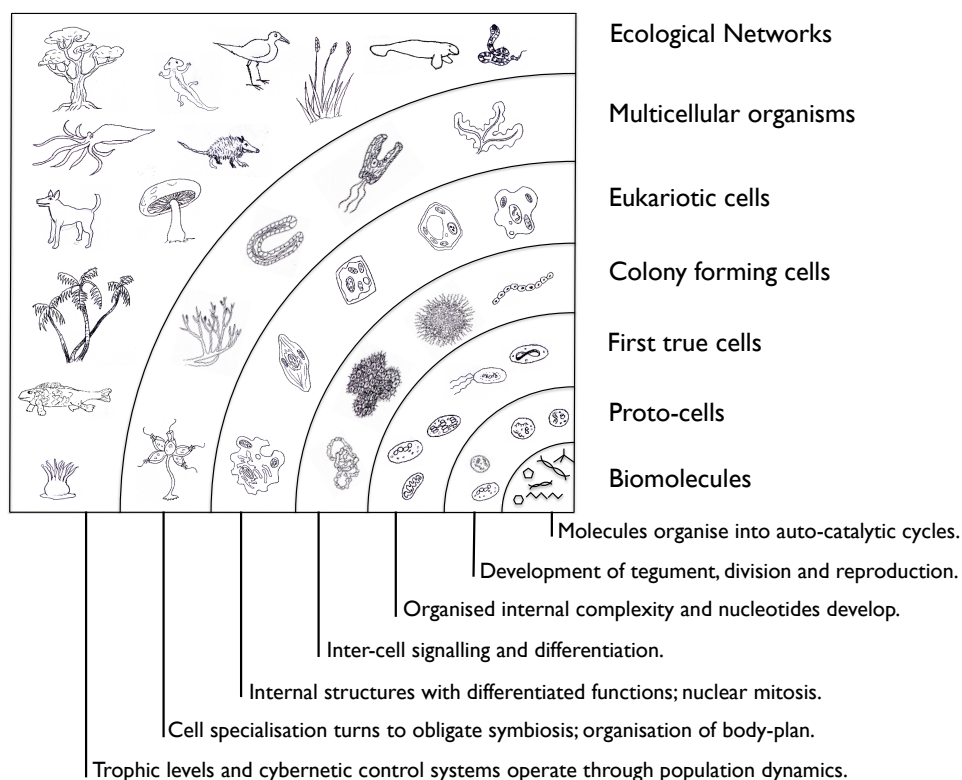
We have argued that life is a dynamic process of filtering and communicating information. The processing of information (computation) occurs in all cases of changing, combining and directing information. Thus computation is a natural, continuous and ubiquitous process (see Denning (2007)). However, it is important to distinguish between (a) universal computing, which can represent any computation in symbols that may be ‘programmed’ and (b) fixed computing in which the hardware and software are interdependent, so that only a narrow range of computational tasks may be performed (this point is discussed by Hopfield (1994)). Life is very much in the latter category (though since the brain is one of its products, this is not universally the case). Complex system computation is now a well established model in behavioral ecol-

ogy, describing many aspects of social organization (reviewed by [Camazine et al. 2001](#)). Other kinds of computation performed by life include information replication, ordering and re-ordering of form and cybernetic system control, each of which will be briefly illustrated below. In each case, computation occurs on a distributed network ([Gershenson, 2010](#)), rather than through the linear Von-Neumann architecture of the familiar digital computer. Whether looking at molecular networks or ecological communities, we see that natural computation is composed of cybernetic feedback loops arranged functionally so that the system gains in persistence. That these loops exist is not a surprise, since any random interconnection of quantities may contain loops and many physical processes do. As control circuits they may generate positive feedback, often leading to quick destruction, or negative feedback leading to stability, and hence more likely to persist in their changing environment. In fact, since control of this kind enhances persistence, natural selection favours cybernetic systems (with negative feedback) above others and we may find this kind of computation practically inevitable. However, a network solely composed of negative feedback fixes on a particular equilibrium, so may be insufficiently flexible to perform the processes of life ([Kauffman, 1993](#)). Since a mix of positive and negative feedback loops can create a dynamic and adaptable system of ‘state-cycles’ in the narrow ‘critical’ regime between catastrophe and order (exemplified by random Boolean networks with high link densities) this has been proposed as an essential feature of living systems by [Kauffman \(1993\)](#) and we now look for evidence of these in significant developments of biological organisation (Figure 2).

### 3.1 Computing through cell-signaling networks

If living is the self-sustaining *coordination* of chemical reactions, does this suggest a coordinating manager? The nucleus was once thought to be the ‘command centre’ of the eukaryotic cell, but observations of cells behaving normally for months after enucleation show that the information processing needed for most activities is cytoplasmic ([Goldman et al., 1973](#)). It would be better to think of the nucleus as the ‘hard disk’ of the cell, since here (for the human) the ‘blueprints’ for at least 47 thousand different proteins ([Orchard et al., 2005](#)) are stored and transcribed, together with editable instructions about when to make them. The ‘algorithms of living’ are run on these proteins which act in ways analogous to transistors and other electronic components, in complex networks, as described by [Butler et al. \(1998\)](#).

The model of cellular information processing as analogue computation (e.g. [Rodbell \(1995\)](#)) was inspired by the cybernetic theory of Norbert Wiener ([1948](#)). In this model, external chemical messages (first messengers) are first ‘discriminated’ (by the receptor) then ‘transduced’ (by a G protein) and finally amplified (by an effector enzyme) to produce an intracellular signal (the second messenger)—a sequence that can



**Fig. 2** Hierarchical self-assembly of complex systems: the increase in computational complexity through the history of life on earth, often associated with a major transition e.g. from prokaryotic to eukaryotic life-forms, or the development of cell-signalling networks or ecological networks. Note that concentric rings indicate expansion of complexity, rather than a chronological sequence: all inner layers exist concurrently at each level.

328 be summarized as perception. This second signal typically initiates a complex sequence of interconnected  
 329 changes which may alter the internal chemistry of the cell, change the response to other first messengers,  
 330 and even selectively alter gene expression (Cairns et al., 1988). Such cascades of molecular response form  
 331 dynamic networks that carry and process information (Lehn, 1990), analogous to artificial neural networks.  
 332 Chemical switches are implemented by the allostery of proteins, especially enzymes, acting as ‘transis-  
 333 tors’ in the network circuitry (Bray, 1995). Furthermore, activated proteins do not simply diffuse to collide  
 334 with their targets. Cytoplasm is a well organized and densely crowded environment in which the reaction  
 335 cascades are localized by ‘scaffold’ proteins, reminiscent of the electronic circuit board. For example, the  
 336 protein kinase enzyme, type II PKA may be fixed to either the plasma membrane, the cytoskeleton, se-  
 337 cretory granules, or the nuclear membrane by anchoring proteins (Scott and Carr, 1992). The effect is not  
 338 only to position this signaling protein close to its intended target but also to determine the local molecular  
 339 environment (context) which may profoundly influence the effect. Such protein networks are built and re-

340 paired following the DNA blueprint, which as we have just noted, may itself be altered by the cytoplasmic  
341 computation. Thus, proteins dynamically send, receive and respond to informational signals in complex  
342 and dynamically changing networks of both negative and positive feedback, which, collectively interacting  
343 with stored DNA-information, form the behavior of the cell and this is readily interpreted as molecular  
344 computation.

### 345 3.2 Replicating information

346 Biological reproduction is an information transfer (communication) phenomenon, from parent(s) as the  
347 transmitter to daughter(s) as the receiver. This biological communication requires a high standard of ac-  
348 curacy, since the information being transmitted is very nearly all functional (Schneider, 2000). Given this  
349 view of reproduction as efficient semantic communication, it was a surprise to realise that the length of  
350 the nuclear genome bears no relation to organism complexity (Gregory, 2001; Valentine, 2003a). Since the  
351 complexity of a system can be defined as the minimum amount of information needed to describe (or repro-  
352 duce) it, one possible reason is that species differ in the amount of error-mitigating repetition their genomes  
353 carry. As well as this, the DNA of almost all organisms harbours a zoo of information parasites (selfish DNA  
354 - Orgel and Crick (1980)) and their remnants, making up a large part of what was historically referred to  
355 as 'junk DNA' when its function was unknown. Transposable elements form the majority of this repetitive  
356 information (Wessler, 2006). It is now thought that many of these 'transposons' originated as endogenised  
357 retro-viruses (Bowen and Jordan, 2002): parasites that have been co-opted into functional symbiosis under  
358 regulation by the host (Veitia and Bottani, 2009). This legacy of non-host information accounts for a large  
359 part of the huge variation in genome size among eukaryotes, where multiple copies of information parasites  
360 are found. However, the relationship between nuclear genome size and organism complexity is still an open  
361 question.

362 Given our understanding of emergence and the formation of functional information from mutual context,  
363 we can see that not all of the functional information is to be found in nuclear DNA. So whilst physically, it  
364 is the genes that are replicated in biological reproduction, context-dependent relationships among them con-  
365 stitutes functional information that is carried along with the replication. Gene regulatory networks (GRNs)  
366 (Davidson and Levin, 2005) are the most significant information complexes to extend beyond nuclear DNA  
367 and are composed of context-dependent relationships among infons, rich in both negative and positive feed-  
368 back. Again, these networks are readily modeled as computational systems (Kravchenko-Balasha et al.,

369 [2012](#)) and their role in determining body-plan through epigenetic phenomena points to a possible correla-  
370 tion between GRN complexity (hence information content) and organism complexity.

### 371 3.3 The eukaryotic revolution

372 Following pioneering work by [Margulis \(1970\)](#), endosymbiosis is the front-running theory explaining the  
373 origin of eukaryotic cells and this well illustrates the increase of function brought about through the creation  
374 of mutual context among infons (level 2 in Figure 1). The advantage of eukarotic cells over prokaryotic is  
375 the specialisation of metabolic, anabolic and reproductive machinery. The component parts collectively  
376 become more efficient by (a) individually concentrating on a smaller task and (b) sharing the products.  
377 The fundamental reason this narrowing of tasks improves effectiveness is that it reduces the information  
378 requirement for performing all necessary tasks. If we think of a cell as a machine performing  $n$  processes;  
379 it needs storage capacity enough to instantiate the algorithms for all  $n$  tasks. prokaryotic cells have rather  
380 limited storage capacity (determined by their AIC), so cannot afford a very sophisticated algorithm for  
381 every task they have to perform - they are limited in effectiveness by their information capacity limit.  
382 When a cell incorporates others, it increases its storage capacity and permits a distribution of tasks among  
383 specialist components, each of which can devote the whole of their limited storage capacity to carrying a  
384 sophisticated and efficient algorithm for a single task. It is also necessary to include the communications  
385 and sharing among the specialist components, so some algorithm space is devoted to this. The exchange  
386 among individual components forms a network of control computation, which on a larger scale constitutes  
387 a complex system (level 3 in Figure 1).

### 388 3.4 Cell types and body-plan complexity

389 Information's role in ordering of form is most apparent in the building of multi-cellular organisms. Cells  
390 come in a large variety of forms, with hierarchical morphotype structure and developmental lineages ([Valen-  
391 tine, 2003a](#)). The number of distinct cell types in a single organism is taken as an indicator of its complex-  
392 ity ([Carroll, 2001](#)) and varies among metazoan phyla from 3 (Myxozoa) to 210 (human) having steadily  
393 increased through evolutionary time ([Valentine et al., 1994](#)). This indicates a gradual accumulation of bio-  
394 logical complexity, and therefore functional information, as life-forms have radiated and cell specialisation  
395 has apparently increased. Despite that, [Hinegardner and Engelberg \(1983\)](#) concluded that "evolution since  
396 the Cambrian appears to have involved few major increases in biological complexity", as [Valentine \(1994\)](#)

397 argued, the basic body-plans of all extant phyla were established by the end of the Cambrian explosion  
398 (520 My ago). The apparent contradiction may be explained by proliferating patterns of gene expression,  
399 rather than the creation of new genes; this being one of the central hypotheses of evolutionary development  
400 biology (see [Valentine, 2003b](#)). Such proliferation of patterns and the consequent radiation of organism-  
401 forms is the result of ordering and re-ordering of functional information. Different cell-types are created by  
402 regulating the expression of different genes in the total genome—simpler organisms suppress the expres-  
403 sion of a higher proportion of their developmental genes than do complex ones ([Davidson, 2001](#)). Thus, the  
404 morphological complexity of an organism is determined by the regulatory machinery which selects genetic  
405 expression during the development of an organism. The number of cell types is one computed ‘output’ of  
406 gene regulatory networks and gives a very rough indication of functional information content. A trend in  
407 modeling body-plan regulatory networks, represents them in a way analogous to artificial neural networks  
408 ([Geard and Wiles, 2005](#)), clearly interpreting morphogenesis as computation. This suggests a means of  
409 quantifying the functional information of body plans by experimentally (*in silico*) examining variants of  
410 formative gene-networks and recording the resulting morphometric diversity.

### 411 3.5 Cybernetic computation by ecological communities

412 Darwin’s metaphor of a ‘tangled bank’ suggests a bewildering complex of interactions among whole or-  
413 ganisms ([Montoya et al., 2006](#)), but natural computation is rarely, if ever, explicit in ecological models.  
414 Information processing in ecological communities is less clear than in cells and organisms because ecosys-  
415 tems usually lack obvious boundaries and their functions are usually considered, not at the system level, but  
416 at the population level, where cybernetic control is not apparent. However, some recent developments pave  
417 the way for this to change; both in describing the information content of communities and in understanding  
418 them as self-regulating complex systems.

419 The study of biodiversity provides a starting point to finding the functional information content at the  
420 ecological level. Using the idea that difference is the basis of information ([Floridi, 2005](#)), diversity (which  
421 by definition counts total difference) becomes a measure of information content. Traditionally, biodiversity  
422 describes the number of different species and perhaps the evenness of their abundances in an ecological  
423 assembly, using metrics inspired by Shannon’s information theory (see [Magurran, 2004](#)). More recently,  
424 broader definitions recognize diversity at every level in the biological hierarchy (table 1), and ecologists may  
425 now refer to genetic and functional diversity as equally necessary for specifying biodiversity ([Lyashevskaya  
426 and Farnsworth, 2012](#)). Ecological communities can be regarded as the vaults of information capital, in



427 the form of molecular structures; networks and pathways; cell types; tissues and organs, whole organisms  
428 and community interactions (Farnsworth et al., 2012). However, ecologists still refer to organizational scale  
429 through informal terms: for example ‘alpha’ and ‘beta’ diversity, which are arbitrarily defined phenomena  
430 of classification (Tuomisto, 2010) over probability distributions (McGill, 2011; Nekola and White, 1999).  
431 Whilst hierarchical nesting of complex systems is explicitly recognised by multi-level modelling in sub-  
432 cellular biology, the strength of formal description this provides has yet to enter ecology (see Faeder, 2011).

433 Descriptive approaches can be developed into conceptual models by changing the focus towards the  
434 network through which organisms interact (as in Norton and Ulanowicz, 1992). The basic components for  
435 such models are available in the special case of predator-prey interactions (e.g. Dunne et al., 2002) (though  
436 models of other material and informational – e.g. genetic – flows are less well developed). Given a network  
437 description such as a food-web, Farnsworth et al. (2012) showed how the functional information approach  
438 may be applied at the ecological level. They systematically dismantled a network model of the Northeast  
439 Atlantic fish community, at each stage measuring its productivity, to find a relationship between complex-  
440 ity and function, which provided a measure of the marginal change in function with network (algorithmic)  
441 information content. Food-webs are but a partial description of ecosystems, which necessarily include  
442 chemical, energy and information flows. Being relatively simpler, microbial networks are more amenable  
443 to this fuller description. The recent development of functional and genetic network models in microbial  
444 ecology (e.g. Zhou et al., 2010) gives us a stepping stone between sub-cellular networks and community  
445 level computation. Significantly, microbial colonies preceded the close association of eukaryotic cells to  
446 form multi-cellular organisms Lepot et al. (2008), yet specialisations among microbe species imply the  
447 same need for self-regulating interactions as is found in organismal physiology. By definition, an isolated  
448 microbial community must be autopoietic and as specialisation among constituent species develops, so must  
449 flows of coordinating information work to compute the community, via complex-system emergence.

### 450 3.6 Information processing as an integrated whole.

451 The computation performed by ecological networks is both broader and narrower than that of a Turing ma-  
452 chine (a system following a sequence of logical operations defined by Turing, 1936). It is broader because  
453 inputs are processed continuously, the outputs are produced continuously and because processing is sen-  
454 sitive to the environment (in a Turing machine, processing is blind to all but the initial inputs until a halt  
455 condition is reached, releasing the output). It is narrower because the computation is equivalent to running  
456 a particular model: a model of the system under control, following the injunction of Conant and Ashby

(1970), that “Every good regulator of a system must be a model of that system”. Interactions between the biotic and abiotic spheres of the ecosystem are regulated by controls on chemical flows, driven by the processing of materials by life *in aggregate*: the sum of individual selection and processing actions amounts to a regulation of the whole ecosystem. Including the flows of nutrients such as nitrogen and phosphorus in ecological network analysis (e.g. Ulanowicz and Baird, 1999) takes us a step closer to the biochemical analogy of within-cell computation (Ulanowicz, 1980). Since molecules continually flow through ecosystems, just as they do in the cell, we can identify the process of constant renewal of ecosystem structure (the network) as autopoiesis, this time referring to all life in aggregate. The phenomenon of constant renewal by recycling material, driven by transforming high to low entropy energy, accumulated over all life on earth, is the foundation of the Gaia hypothesis (Lovelock and Margulis, 1974). The total of global ecological processes may be interpreted as a network computer, whose input is the physical and chemical environment of the planet and the output is a computed adjustment of these to maintain equilibrium. Seen this way, life is a computer running a model of itself in order to control its interior state so as to perpetuate itself in a changeable environment. This view, which goes beyond cybernetic self-regulation to reveal autopoietic computation, is closely allied to a growing thermodynamic understanding of living processes in which energy accountancy is integrated with informational interpretations (e.g. in Smith, 2008). For example, the accumulation of hierarchical complexity, so characteristic of life, has been demonstrated to follow from thermodynamic efficiency (Wicken, 1979; Annala and Annala, 2008; Annala and Kuismanen, 2009) as has the tendency for hierarchical complex structures to regulate their internal and external environments through information processing (Kaila and Annala, 2008; Karnani and Annala, 2009).

#### 4 Implications

The information perspective shows life to be (a) continuous with the abiotic universe and (b) the consequence of a spontaneous increase in complexity through repeated combination of formative patterns such that they give context and thence function to one-another. Chemistry is the result of this process at the atomic scale and life is a branch of chemistry that is especially rich in opportunities for functional combinations. The processes of life are chemical processes, so our ‘life is information’ remains compatible with Kornberg’s ‘life is chemistry’, but goes deeper by highlighting the informational basis of the chemistry of life. Our perspective also emphasises the idea that the whole of life at all scales has a role in reproducing life. Considering life as information processing (computation) where the subject of computation is life, we are faced with a ‘program’ running on itself, the function of which is to output itself. Such recursion is

---

487 familiar and much exploited in computer science. It highlights the fact that for life, there is no distinction  
488 between the 'machine' and the program - both are information; they are the same information, ordering and  
489 re-ordering matter and energy so as to persist. It would not be right to think of life as a biochemical structure  
490 on which a program is run, because life is the program and the biochemical structure is its embodiment.  
491 This is why we say that information is not just in DNA, but is in the whole biological system. The idea that  
492 'life is information processing' brings reductionists and synthesists closer together as it shows life to super-  
493 vene on chemistry strictly according to information content, but to also possess strictly emergent properties  
494 (at several levels) arising from the *functions* of the embodied information. Now that functional information  
495 content can be quantified at every level of life, we anticipate its use in further deepening our understanding  
496 of life and its place in the physical universe.

#### 497 Acknowledgments

498 This work was enhanced by very thoughtful and creative reviews by anonymous referees. It was supported  
499 by a Science Technology Research and Innovation for the Environment grant from the Environmental Pro-  
500 tection Agency of the Republic of Ireland: 2007-PhD-SD-3. C.G. was partially supported by SNI member-  
501 ship 47907 of CONACyT, Mexico.

502 **References**

- 503 Adami, C., Ofria, C., and Collier, T. C. (2000). Evolution of biological complexity. *Proc Natl Acad Sci U*  
504 *S A*, 97(9):4463–4468.
- 505 Annala, A. and Annala, E. (2008). Why did life emerge? *International Journal of Astrobiology*, 7(3-4):293–  
506 300.
- 507 Annala, A. and Kuismanen, E. (2009). Natural hierarchy emerges from energy dispersal. *Biosystems*,  
508 95(3):227–233.
- 509 Balleza, E., Alvarez-Buylla, E. R., Chaos, A., Kauffman, S., Shmulevich, I., and Aldana, M. (2008). Critical  
510 dynamics in genetic regulatory networks: Examples from four kingdoms. *Plos One*, 3(6):e2456.
- 511 Bates, M. (2005). Information and knowledge: an evolutionary framework for information science. *Infor-*  
512 *mation Research*, 10(4):paper 239.
- 513 Bateson, G. (1972). Form, substance, and difference. In Bateson, G., editor, *Steps to an Ecology of Mind*,  
514 pages 448–466. University of Chicago Press.
- 515 Bitbol, M. and Luisi, P. (2004). Autopoiesis with or without cognition: defining life at its edge. *J Royal Soc*  
516 *Interface*, 1(1):99–107.
- 517 Bowen, N. and Jordan, I. (2002). Transposable elements and the evolution of eukaryotic complexity. *Curr*  
518 *Issues Mol Biol*, 4:65–76.
- 519 Bray, D. (1995). Protein molecules as computational elements in living cells. *Nature*, 376(6538):307–312.
- 520 Bray, D. (2009). *Wetware: A computer in every living cell*. Yale University Press, New Haven, CT. USA.
- 521 Butler, M. H., Paton, R. C., and Leng, P. H. (1998). *Information processing in tissues and cells*, chapter  
522 Information processing in computational tissues, pages 177–184. Plenum Press, New York.
- 523 Cairns, J., Overbaugh, J., and Miller, S. (1988). The origin of mutants. *Nature*, 335:142–145.
- 524 Camazine, S., Deneubourg, J. L., Franks, N., Sneyd, J., Theraulaz, G., and Bonabeau, E. (2001). *Self-*  
525 *Organization in Biological Systems*. Princeton University Press, Princeton, NJ. USA.
- 526 Carroll, S. (2001). Chance and necessity: the evolution of morphological complexity and diversity. *Nature*,  
527 409(6823):1102–1109.
- 528 Chaitin, G. (1990). *Information, Randomness and Incompleteness - Papers on Algorithmic Information*  
529 *Theory*, volume 8 of *Series in Computer Science*. World Scientific, Singapore, 2nd edition.
- 530 Conant, R. and Ashby, W. (1970). Every good regulator of a system must be a model of that system.  
531 *International Journal of Systems Science*, 1(2):89–97.

- 532 Cowling, R., Knight, A., Faith, D., Ferrier, S., Lombard, A., Driver, A., Rouget, M., Maze, K., and Desmet,  
533 P. (2004). Nature conservation requires more than a passion for species. *Conserv Biol*, 18(6):1674–1676.
- 534 Cummins, R. (1975). Functional analysis. *J. Philos.*, 72(20):741–765.
- 535 Curtis, T., Sloan, W., and Scannell, J. (2002). Estimating prokaryotic diversity and its limits. *Proc Natl*  
536 *Acad Sci U S A*, 99(16):10494–10499.
- 537 Davidson, E. H. (2001). *Genomic regulatory systems: Development and evolution*. Academic Press, San  
538 Diego, USA.
- 539 Davidson, E. H. and Levin, M. (2005). Gene regulatory networks. *Proc Nat Acad Sci USA*, 102(14):4935.
- 540 Denning, P. J. (2007). Computing is a natural science. *Communications of the ACM*, 50(7):13–18.
- 541 Dunne, J., Williams, R., and Martinez, N. (2002). Food-web structure and network theory: The role of  
542 connectance and size. *Proc Natl Acad Sci U S A*, 99(20):12917–12922.
- 543 Faeder, J. R. (2011). Toward a comprehensive language for biological systems. *BMC Biol*, 9:68.
- 544 Farnsworth, K., Lyashevskaya, O., and Fung, T. (2012). Functional complexity: The source of value in biodi-  
545 versity. *Ecol Complex*, 11:46–52.
- 546 Favareau, D., editor (2009). *Essential Readings in Biosemiotics: Anthology and Commentary*. Springer,  
547 Berlin.
- 548 Floridi, L. (2003). Information. In Floridi, L., editor, *The Blackwell Guide to the Philosophy of Computing*  
549 *and Information*, pages 40–61. Blackwell Publishing Ltd.
- 550 Floridi, L. (2005). Is semantic information meaningful data? *Philosophy and Phenomenological Research*,  
551 70(2):351–370.
- 552 Galtin, L. L. (1972). *Information Theory and the Living System*. Columbia University Press, New York.
- 553 Geard, N. and Wiles, J. (2005). A gene network model for developing cell lineages. *Artif Life*, 11:249–267.
- 554 Gell-Mann, M. and Lloyd, S. (1996). Information measures, effective complexity, and total information.  
555 *Complexity*, 2(1):44–52.
- 556 Gell-Mann, M. and Lloyd, S. (2003). Effective complexity. In Gell-Mann, M. and Tsallis, C., editors,  
557 *Nonextensive Entropy - Interdisciplinary Applications*. Oxford University Press.
- 558 Gershenson, C. (2010). Computing networks: A general framework to contrast neural and swarm cogni-  
559 tions. *Paladyn, Journal of Behavioral Robotics*, 1(2):147–153.
- 560 Gershenson, C. and Fernández, N. (2012). Complexity and information: Measuring emergence, self-  
561 organization, and homeostasis at multiple scales. *Complexity*, Early View.

- 562 Goldman, R., Pollack, R., and Hopkins, N. (1973). Preservation of normal behavior by enucleated cells in  
563 culture. *Proc Nat Acad Sci USA*, 70:750–754.
- 564 Gregory, T. (2001). Coincidence, coevolution, or causation? DNA content, cell size, and the C-value  
565 enigma. *Biological Reviews*, 76(1):65–101.
- 566 Griffiths, P. E. (1993). Functional analysis and proper functions. *British J. Philos. Sci.*, 44:409–422.
- 567 Hinegardner, R. and Engelberg, J. (1983). Biological complexity. *J Theor Biol*, 104:7–20.
- 568 Hopfield, J. J. (1994). Physics, computation, and why biology looks so different. *J Theor Biol*, 171:53–60.
- 569 Jiang, Y. and Xu, C. (2010). The calculation of information and organismal complexity. *Biol Direct*, 5:59.
- 570 Kaila, V. R. I. and Annala, A. (2008). Natural selection for least action. *Proceedings of the Royal Society*  
571 *A-Mathematical Physical and Engineering Sciences*, 464(2099):3055–3070.
- 572 Karnani, M. and Annala, A. (2009). Gaia again. *Biosystems*, 95(1):82–87.
- 573 Kauffman, S. A. (1993). *Origins of Order: Self-Organization and Selection in Evolution*. Oxford University  
574 Press, Oxford, UK.
- 575 Kohl, P., Crampin, E. J., Quinn, T. A., and Noble, D. (2010). Systems biology: An approach. *Clin Pharma-*  
576 *col Ther*, 88(1):25–33.
- 577 Kornberg, A. (1991). Understanding life as chemistry. *Clin Chem*, 37(11):1895–1899.
- 578 Kravchenko-Balasha, N., Levitzki, A., Goldstein, A., Rotter, V., Gross, A., Remacle, F., and Levine, R. D.  
579 (2012). On a fundamental structure of gene networks in living cells. *Proc Natl Acad Sci U S A*,  
580 109(12):4702–7.
- 581 Lee, K. (2004). There is biodiversity and biodiversity. In Oksanen, M. and Pietarinen, J., editors, *Philosophy*  
582 *and Biodiversity*, pages 152–171. Cambridge University Press, Cambridge, UK.
- 583 Lehn, J.-M. (1990). Perspectives in supramolecular chemistry—from molecular recognition towards molec-  
584 ular information processing and self-organization. *Angewandte Chemie International Edition in English*,  
585 29(11):1304–1319.
- 586 Lepot, K., Benzerara, K., Brown, G., and P, P. (2008). Microbially influenced formation of 2,724-million-  
587 year-old stromatolites. *Nat. Geosci.*, 1:118–121.
- 588 Li, M. and Vitányi, P. M. B. (2008). *An introduction to Kolmogorov complexity and its applications*.  
589 Springer, 3rd edition.
- 590 Lorenz, D. M., Jeng, A., and Deem, M. W. (2011). The emergence of modularity in biological systems.  
591 *Physics of Life Reviews*, 8(2):129 – 160.

- 592 Lovelock, J. E. and Margulis, L. (1974). Atmospheric homeostasis by and for the biosphere: The Gaia  
593 hypothesis. *Tellus*, 26(1):2–10.
- 594 Lyashevskaya, O. and Farnsworth, K. D. (2012). How many dimensions of biodiversity do we need? *Ecolog-  
595 ical Indicators*, 18:485–492.
- 596 MacKay, D. M. (1969). *Information, Mechanism and Meaning*. MIT Press, Cambridge, MA, USA.
- 597 Magurran, A. (2004). *Measuring Biological Diversity*. Blackwell Publishing.
- 598 Margulis, L. (1970). *Origin of Eukaryotic Cells*. Yale University Press, New Haven, CT. USA.
- 599 Maturana, H. and Varela, F. J. (1980). *Autopoiesis and Cognition: the Realization of the Living*. D. Reidel  
600 Publishing Company, Dordrecht, NL. Translation of original: De Maquinas y seres vivos. Universitaria  
601 Santiago.
- 602 Maus, C., Rybacki, S., and Uhrmacher, A. M. (2011). Rule-based multi-level modeling of cell biological  
603 systems. *BMC Syst Biol*, 5:166.
- 604 McAllister, J. (2003). Effective complexity as a measure of information content. *Philos Sci*, 70(2):302–307.
- 605 McGill, B. J. (2011). Linking biodiversity patterns by autocorrelated random sampling. *Am J Bot*,  
606 98(3):481–502.
- 607 Menconi, G. (2005). Sublinear growth of information in dna sequences. *Bull. Math. Biol.*, 67(4):737–759.
- 608 Montoya, J., Pimm, S. L., and Solé, R. V. (2006). Ecological networks and their fragility. *Nature*,  
609 442(7100):259–264.
- 610 Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., and Worm, B. (2011). How many species are there  
611 on earth and in the ocean? *PLoS Biol*, 9(8):e1001127.
- 612 Morowitz, H. J. (1992). *Beginnings of Cellular Life*. Yale University Press, New Haven, CT. USA.
- 613 Neander, K. (1991). Functions as selected effects: A conceptual analysts defense. *Philos. Sci.*, 58(2):168–  
614 184.
- 615 Neander, K. (2011). *Routledge Encyclopedia of Philosophy (Online)*. Routledge.
- 616 Nekola, J. and White, P. (1999). The distance decay of similarity in biogeography and ecology. *J Biogeogr*,  
617 26(4):867–878.
- 618 Norton, B. and Ulanowicz, R. (1992). Scale and biodiversity policy - a hierarchical approach. *Ambio*,  
619 21(3):244–249.
- 620 Orchard, S., Hermjakob, H., and Apweiler, R. (2005). Annotating the human proteome. *Mol Cell Pro-  
621 teomics*, 4(4):435–40.
- 622 Orgel, L. and Crick, F. (1980). Selfish DNA: The ultimate parasite. *Nature*, 284:604–607.

- 623 Prigogine, I. (1977). *Self-Organization in Non-Equilibrium Systems*. Wiley, New York.
- 624 Prigogine, I. and Stengers, I. (1984). *Order out of Chaos: Man's new dialogue with nature*. Flamingo.  
625 Collins Publishing Group., London.
- 626 Robertson, M. and Joyce, G. (2010). The origins of the rna world. *Cold Spring Harbour Perspectives In*  
627 *Biology*.
- 628 Rodbell, M. (1995). Signal transduction: evolution of an idea. *Biosci Rep.*, 15:117–133.
- 629 Salthe, S. (1985). *Evolving Hierarchical Systems: Their Structure and Representation*. Columbia University  
630 Press.
- 631 Schneider, T. D. (2000). Evolution of biological information. *Nucleic Acids Res*, 28:2794–2799.
- 632 Schrödinger, E. (1944). What is Life? The physical aspects of the living cell.  
633 <http://home.att.net/~p.caimi/schrodinger.html>.
- 634 Scott, J. and Carr, W. (1992). Subcellular localization of the type II cAMP-dependent protein kinase.  
635 *Physiology*, 7:143–148.
- 636 Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3,4):379–  
637 423,623–656.
- 638 Smith, E. (2008). Thermodynamics of natural selection i: Energy flow and the limits on organization. *J*  
639 *Theor Biol*, 252(2):185–197.
- 640 Smith, E. and Morowitz, H. J. (2004). Universality in intermediary metabolism. *Proc Nat Acad Sci USA*,  
641 101(36):13168–13173.
- 642 Szostak, J. W. (2003). Functional information: Molecular messages. *Nature*, 423(6941):689–689.
- 643 Tuomisto, H. (2010). A diversity of beta diversities: straightening up a concept gone awry. part 1. defining  
644 beta diversity as a function of alpha and gamma diversity. *Ecography*, 33(1):2–22.
- 645 Turing, A. (1936). On computable numbers, with an application to the entscheidungs problem. *Proceedings*  
646 *of the London Mathematical Society*, 42:230–265.
- 647 Turing, A. (1952). The chemical basis for morphogenesis. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.*,  
648 237:37–72.
- 649 Ulanowicz, R. (1980). An hypothesis on the development of natural communities. *J Theor Biol*, 85:223–  
650 245.
- 651 Ulanowicz, R. and Baird, D. (1999). Nutrient controls on ecosystem dynamics: The chesapeake mesohaline  
652 community. *J Mar Syst*, 19:159–172.



- 653 Valentine, J. (1994). Late precambrian bilaterians: grades and clades. *Proc Natl Acad Sci U S A*,  
654 91(15):6751–6757.
- 655 Valentine, J. (2003a). Architectures of biological complexity. *Integrative and comparative biology*,  
656 43(1):99–103.
- 657 Valentine, J. (2003b). Cell types, cell type numbers, and body plan complexity. In Hall, B. and Olson, W.,  
658 editors, *Keywords and Concepts in Evolutionary Developmental Biology*, pages 35–43. Harvard Univer-  
659 sity Press, Cambridge, MA, USA.
- 660 Valentine, J., Collins, A., and Meyer, C. (1994). Morphological complexity increase in metazoans. *Paleo-*  
661 *biology*, 20(2):131–142.
- 662 Veitia, R. A. and Bottani, S. (2009). Whole genome duplications and a ‘function’ for junk DNA? Facts and  
663 hypotheses. *Plos ONE*, 4(12):e8201.
- 664 von Foerster, H. (1960). On self-organizing systems and their environments. In Yovits, M. and Cameron,  
665 S., editors, *Self-organizing systems*. Pergamon Press, Oxford, UK.
- 666 Wessler, S. R. (2006). Transposable elements and the evolution of eukaryotic genomes. *Proc Natl Acad Sci*  
667 *U S A*, 103(47):17600–17601.
- 668 Wicken, J. S. (1979). The generation of complexity in evolution: A thermodynamic and information-  
669 theoretical discussion. *J Theor Biol*, 77:349–365.
- 670 Wiener, N. (1948). *Cybernetics; or, Control and Communication in the Animal and the Machine*. Wiley  
671 and Sons, New York.
- 672 Yockey, H., Platzman, R., and Quastler, H., editors (1958). *Symposium on Information Theory in Biology*  
673 *(1956 : Gatlinburg, Tenn.)*. Pergamon Press, New York.
- 674 Zepik, H., Blochliger, E., and Luisi, P. (2001). A chemical model of homeostasis. *Angewandte Chemie-*  
675 *International Edition*, 40(1):199–202.
- 676 Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., and Zhi, X. (2010). Functional molecular ecological networks.  
677 *MBio*, 1(4):e00169–00110.

<b>Organization Level</b>	<b>Interactions</b>
life as a whole	global bio-geochemical networks
ecological communities	interspecific material and energy flows
populations - species	gene-flow, dispersal, evolution
multi-cellular organisms	organism societies + interspecific, e.g. parasitism
tissues, organs and organ systems	cellular communication and organ function
cells	specialisation and ontogeny: e.g. immune system
sub-cellular structures	catabolic autopoietic processes
molecular networks	metabolic and information processing
DNA sequences: codons to genes	coding and expression control
molecular surfaces	lock and key - enzymes

**Table 1** A ten-level hierarchy of biocomplexity. Left column names the level of organization and right column gives examples of the complex interactions and processes that take place at that level, contributing to biocomplexity. Complexity is also added by interactions among levels, both upwards and downwards, producing feedback circuits. Interactions at every level and among levels constitute information processing. (adapted from [Farnsworth et al. \(2012\)](#))