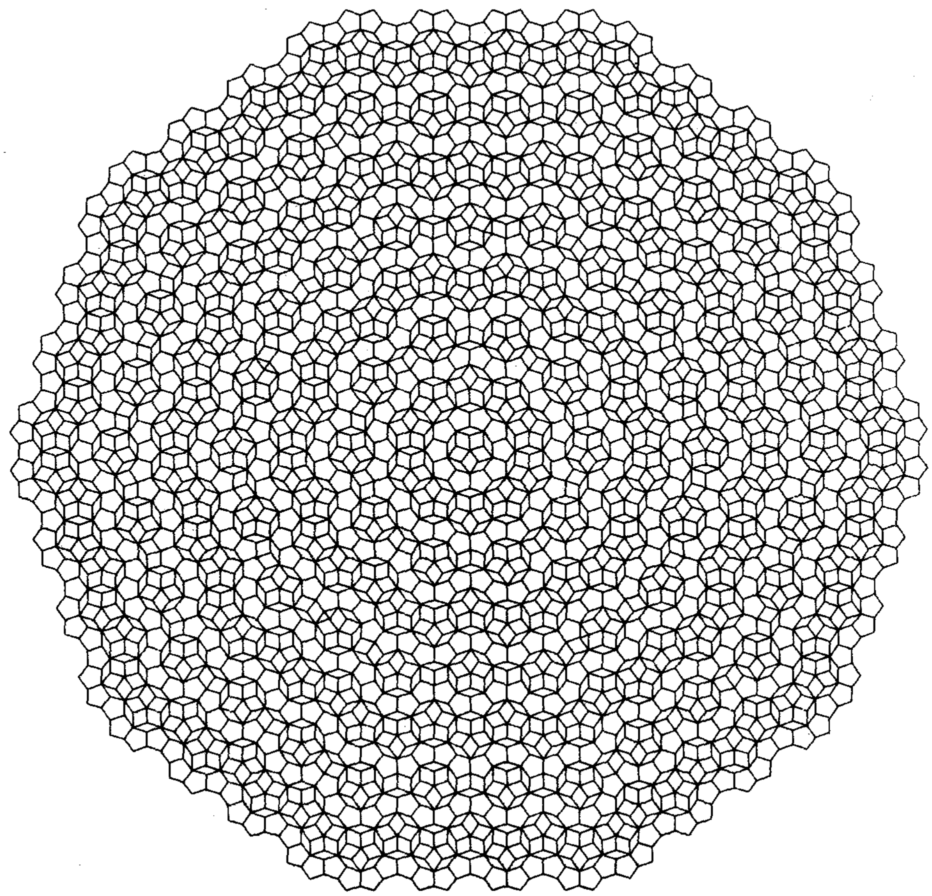
# Symmetry: Culture and Science

Symmetry and Information

The Quarterly of the
International Society for the
Interdisciplinary Study of Symmetry
(ISIS-Symmetry)

Editors:
György Darvas and Dénes Nagy

Volume 7, Number 3, 1996



*Symmetry: Culture and Science*
*Vol. 7, No. 3, 247-256, 1996*

# INFORMATION ORIGINATES IN SYMMETRY BREAKING

John Collier

*Address*: Department of Philosophy, University of Newcastle, Callaghan, NSW 2308, Australia; E-mail: pljdc@alinga.newcasle.edu.au

**Abstract:** *We find symmetry attractive. It interests us. Symmetry is often an indicator of the deep structure of things, whether they be natural phenomena, or the creations of artists. For example, the most fundamental conservation laws of physics are all based in symmetry. Similarly, the symmetries found in religious art throughout the world are intended to draw attention to deep spiritual truths. Not only do we find symmetry pleasing, but its discovery is often also surprising and illuminating as well. For these reasons, we are inclined to think that symmetries are informative, and that symmetries contain information. On the other hand, symmetries represent a kind of invariance under transformation. Such invariance implies that symmetrical things contain redundancies. Redundancy, in turn, implies that the information content of a symmetrical structure or configuration is less than that of a similar nonsymmetrical structure. Symmetry, then, entails a reduction in information content. These considerations present us with somewhat of a paradox: On the one hand, many symmetries that we find in the world are surprising, and surprise indicates informativeness. On the other hand, the surprise value of information arises because it presents us with the unexpected or improbable, but symmetries, far from creating the unexpected, ensure that the known can be extended through invariant transformations. How can this paradox be resolved?*

## 1. THE INFORMATION CONTENT OF BELIEFS

In part, the resolution of the paradox involves disentangling the information content of beliefs about a symmetry from the information in symmetrical objects themselves. Although the information content of a symmetrical configuration is relatively low, knowledge that some system is symmetrical reduces what we need to know about the system by eliminating possibilities that would be permitted if the system were not symmetrical. This reduction of required information is greater the more pervasive the symmetry. The relatively low information content resulting from symmetries is reflected in the high epistemic value of knowledge of these symmetries.

The information content of a belief that there is a symmetry in some structure or configuration is a function of the reduction in the number of possible configurations resulting from the elimination of all the ones that are not symmetrical. The more the supposed symmetries reduce the number of possibilities, the greater the information content of the belief in the symmetries. A convenient way to represent the situation was developed by Carnap and Bar-Hillel (Bar-Hillel, 1964) They used the resources of inductive logic to define the information content of a statement in a given language in terms of the possible states it rules out from a complete ensemble of states, or *state space*. For "technical reasons", largely derived from logical empiricist views of language, they calculated the states ruled out as a number of *state descriptions*. A state description is a conjunction of atomic statements assigning each primitive monadic predicate or its negation (but never both) to each individual constant of the language. The information content of a statement is thus relative to a language. This presents no problem as long as the language is powerful enough to represent all possible states in the state space. Evidence, in the form of *observation statements*, contains information in virtue of the class of state descriptions the evidence rules out[1]. Information content, then, is inversely related to probability, as intuition would suggest.

It turns out, though, that our pre-systematic intuitions confuse two different measures of information content, both of which have plausible but incompatible properties. The first measure of the information content of statement $s$ is called the *content measure*, *cont*($s$). It is defined as the complement of the *a priori* probability that $s$ is true:

$$cont(s) = 1 - prob(s)$$

The *cont* measure correctly represents the complementary character of information and likely configurations, but it has a serious problem. *Cont* fails the *additivity condition*, according to which the combined information content of two inductively independent statements[2] should be the sum of their individual information contents (Bar-Hillel, 1964: 302). This condition is required in the present context, since we are interested in the information difference between a state description without symmetry and the state description with symmetry. The *cont* measure also fails some natural assumptions about conditional information. These problems motivated the introduction of another measure, called the *information measure*, *inf*($s$):

$$inf(s) = \log_2 \frac{1}{1 - cont(s)} = -\log_2 prob(s)$$

The value of this measure is in bits. Although *inf* satisfies additivity and conditionalization requirements, it has a property that some people find counter-intuitive. If some evidence $e$ is negatively relevant to a statement $s$, then the information measure of $s$ conditional on $e$ will be greater than the absolute information measure of $s$. This violates a common intuition that the information of $s$ given $e$ must be less than or equal to the absolute information of $s$. The content measure, *cont*($s$), does satisfy this intuition (Bar-Hillel, 1964: 306-7). Personally, I do not share this widespread intuition, since it requires effort to correct the inference based on $e$ that $s$ is less likely. The issue

[1]Carnap and Bar-Hillel, as was the custom at the time, assumed that observation statements can be connected to experience unambiguously. This assumption turns out to be problematic; I discuss the problems in (Collier, 1990). While no resolution is widely accepted today, the problem does not seriously effect the general principle proposed by Carnap and Bar-Hillel.

[2]Inductive independence means that the conditional probability of each statement given the other is the same as its initial probability.

requires further analysis, but is not relevant in what follows.

On the *inf* measure (and the *cont* measure as well), if there is no knowledge of the state of the system, all state descriptions are equally likely, and *inf*($s$) takes on its lowest possible value. Any evidence which disturbs this equiprobability, including evidence of symmetries, increases the information we have of about the system. The information involved will be the difference between the entropy of the ensemble of equiprobable states and that of the ensemble of states permitted by the evidence. The latter entropy represents the information content of the symmetries in the system, assuming that the only evidence we have is about symmetries. In other words, knowledge of symmetries within a system produces information about the possible configurations ruled out by the symmetries. This information will be greater the less the information in the symmetries, this information being less the more pervasive the symmetries in question.

It would be hasty, however, to think that this completely resolves the paradox. Any new knowledge has much the same effect as knowledge of symmetry. Consider an experiment (a roll of a die, for example), with possible outcomes $O_1$, ... $O_n$, in which we discover $O_i$ is the case. $O_i$ reduces the possibilities in much the same way that learning of a symmetry does. The account does not tell us in any way what is especially interesting about symmetries. Symmetries are interesting over and above their interest as the outcomes of experiments. Another difference is that, while the outcome of an arbitrary experiment might be quite specific, and hence have a relatively high information content, symmetries are always of relatively low intrinsic information content. We still need to explain the interest of symmetries, despite their relatively low information content. I will offer an answer that distinguishes between surface and deep symmetries, and argues that deep symmetries are interesting because of the way that information originates in symmetry breaking. I will first look at the mathematics of information and symmetry, and then at some specific ways in which information originates in thermodynamic, cosmological, biological and perceptual processes. I will conclude with some speculations on why we find symmetry interesting and attractive.

## 2. INFORMATION AND REDUNDANCY

There are three (almost) equivalent approaches to mathematical information theory: combinatorial, probabilistic and computational (Kolmogorov, 1965). On all three approaches, redundancy is a reduction in information. The combinatorial and probabilistic approaches are essentially statistical, and apply most aptly to ensembles of similar systems, although they can be adapted to deal with individual systems with statistical properties (Penrose, 1970; Collier, 1990). The computational approach is more general, since it can apply to any system showing order of some sort, but it is most usefully applied to individual, non-statistical systems (Collier, 1990).

On the combinatorial approach, information depends on frequencies of microstates making up a macrostate within an ensemble, or the frequencies of identical components within an overall configuration:

$$I_{Comb} = \sum_s freq(s) \log_2 \frac{1}{freq(s)}$$

The probabilistic account is similar, except that probabilities rather than frequencies are used:

$$I_{Prob} = \sum_s prob(s) \log_2 \frac{1}{prob(s)}$$

As in the combinatorial account, the probabilities can refer to probabilities in an ensemble of configurations, or to those of elements in a single configuration, though in this case the former interpretation is more natural. These definitions of information can be taken to be entirely abstract, without regard to their physical embodiment, but when they refer to physical systems, the possible states are called *complexions*. In what follows, I will assume that information is taken to be physically embodied. In both the combinatorial or probabilistic versions, if the frequencies or probabilities are equal, then the equations reduce to:

$$I_{Equil} = \log_2 \frac{1}{N}$$

where $N$ is the number of possible complexions. Under this condition, average information content per complexion is maximised[3]. It is worth noting that this is a condition of strong symmetry, since all complexions have equal information value, and are statistically identical. It must be remembered, however, that the symmetry is statistical only, and shows up only at the macroscopic level, in which averages are taken over ensembles or sequences over time. The symmetry involved is the relatively trivial one that characterises all systems in equilibrium: equilibrium states are statistically equivalent, and non-statistical information is eliminated by the averaging process. There is no reason (nor is any assumed) why the complexions should show any non-statistical symmetries. A further point to note is that the average information per complexion in (5) is the negative value of the entropy of the system. This relation is known as the *negentropy principle of information* (Brillouin, 1962; Penrose, 1970).

The computational, or algorithmic approach (Kolmogorov, 1965; 1968; Chaitin, 1975; Li and Vitànyi, 1990) is motivated quite differently. The fundamental hypothesis of algorithmic information theory is that the information content of something is the length of the shortest (self-delimiting) program in binary form that can produce it:

$$I_{Alg}(s) = length(s)$$

The idea behind this hypothesis is that a thing can be specified by making a series of binary distinctions (Spencer Brown, 1972). The minimum number of distinctions required is its information content in bits[4]. The algorithmic definition of information

[3]In the purely abstract representation, without regard to physical embodiment, the value given in equation (5) is the maximal possible average information per message, assuming perfect coding. Unlike physical entropy, it can spontaneously reduce when passed through a passive filter.

[4]The logic of distinctions can be shown to be equivalent to Boolean algebra (Banaschewski, 1977). This reflects the relation between information and the minimal number of distinctions, where each bit of information can be thought of as representing the answer 'true' or 'false' to a set of well chosen questions, much like a perfect game

content is equivalent to the combinatorial and probabilistic information contents except for an additive constant representing computational overhead (Kolmogorov, 1968) that can be made arbitrarily small (Chaitin, 1975). The basic concepts of information theory can be defined without recourse to probability theory, and are applicable to individual cases (Kolmogorov, 1968). Furthermore, the relations between information and probability allow probability theory to be based on algorithmic information theory.

On each of these accounts, redundancy is either a direct repetition of information, or else a correlation within the structure that constrains the amount of information that it can hold. The simplest sort of redundancy is simple repetition of elements. Another form of redundancy occurs when each element constrains the next possible element (Markov redundancy). Higher order redundancies are also possible, in which correlations occur over sequences of N units that cannot be found in fewer than N units. These show up in the presence of deep symmetries. The basic mathematical theory of information, including levels of redundancy, was developed by Shannon (Shannon and Weaver, 1949), but the notion of depth is much more recent. The only suitable formal definition at present is *logical depth* as defined by Charles Bennett (Li and Vitànyi, 1990). It is a measure of the computation time (or number of steps) required to produce the surface structure of a sequence from its most compressed form. Simple repetitive redundancy is relatively easy to compute, but redundancies relying on long sequences of elements, or on mutually constraining elements spread broadly over a sequence, are much harder to compute from their maximally compressed form.

## 3. SYMMETRY DYNAMICS

A symmetry is fundamentally an invariance of a configuration of elements under a group of automorphic transformations (Weyl, 1952). An automorphic transformation is one that maps the elements of a system onto elements of the system such that each element is mapped onto exactly one element by the transformation. Symmetry can be defined as an automorphism, and the degree of symmetry can be completely characterised by the set of transformations that bring the system or configuration into coincidence with itself. (Aleksandrov et al, 1969: Vol. III, pp. 271-2). The identity transformation exists for any system. If the identity transformation is the only automorphism, the system is unsymmetrical. Other automorphisms produce non-trivial transformations, indicating a non-trivial invariance. This invariance implies that the configuration is redundant, because it implies correlations within the system or configuration. The complementary mathematical character of the relation between information and symmetry suggests that information might originate in symmetry reduction, but it is possible that information might have other sources than symmetry breaking. I will argue against this possibility on the grounds of the close tie between equilibrium and symmetry (Weyl, 1952: 25) mentioned above.

With these formal definitions of information and symmetry, we can examine where information originates in the universe. If all nature were entirely lawful, then nature would be fully symmetrical, reflecting the dynamical symmetry of the laws (Weyl, 1952). In order to account for non-symmetries, we need contingency, which arises either from initial conditions, or else from symmetry breaking processes. On the "no boundary conditions" cosmology favoured by many modern cosmologists, there is no

---

of "20 questions".

information in the initial conditions (they are entirely symmetrical) (Layzer, 1990), so all information must arise through symmetry breaking. Any cosmology that does not make this assumption is necessarily incomplete, since any deviation from statistical equivalence is in need of further explanation. Since the no boundary conditions condition implies that all complexions are statistically equivalent in the original state of the universe, the initial cosmological condition is one of macroscopic equilibrium. Unless this is so, we are faced with an infinite regress of cosmological explanations. Subsequent symmetry breaking that produces information occurs in processes in which microscopic fluctuations are promoted to macroscopically detectable structure.

There are two types of symmetry breaking that occur in the early universe. The first is the differentiation of matter and radiation. This occurs when the rate of equilibration between matter and energy (the relaxation rate) falls behind the expansion rate of the universe. This leads to the formation of branch systems that require more information to describe than either the original statistically uniform condition or the local equilibrium condition that immediately precedes the phase separation of matter and radiation. Layzer notes that there is not only structural order in the universe, but also chemical order. The origin of chemical order is a major focus of Steven Weinberg's book, *The First Three Minutes* (Weinberg, 1977), which deals primarily with the formation of matter. As Weinberg points out, the first 1/100 sec is hidden to us because the high energies involved are of the same order as those of nucleons (hadrons in general) (over 100,000 million degrees Kelvin). But the strong force, governing nucleon interactions, has a constant analogous to the fine structure constant of the weak force whose value is close to unity (compared with 1/137 for the fine structure constant), so nucleons interact very strongly. So strongly, in fact, that it is not clear it is possible to separate their components, even if there are some elementary components (varieties of quarks). There is no clear sense in which there is a definite number of particles under these conditions.

At lower temperatures, we get separate particles, but they are in thermal equilibrium until roughly 10,000 million degrees Kelvin (1.09 seconds duration), when neutrinos are no longer in thermal equilibrium. As temperature and density decrease, they work synergistically to lower the rate that fluctuations return to equilibrium. The rate of interactions among particles that causes relaxation to equilibrium depends on the mass density directly, whereas the rate of expansion depends on the square root of mass density, so there is a point at which the rate of expansion exceeds the relaxation rate, and we start to get branch systems forming as the universe goes out of thermal equilibrium. Interestingly, after this time the maximum possible entropy rises faster than the actual entropy, and order and disorder increase together. This permits information to increase at the same time as entropy. Before that time, there is no spatial entropy gradient, and even though global entropy increases, the local conditions are near to equilibrium, since the size of fluctuations is greater than the entropy gradient, and the relaxation time is short enough that fluctuations decay before they can stabilise. The universe was effectively completely symmetrical at the macroscopic scale because it was at equilibrium, and the microstates were statistically equivalent, since they were initially statistically equivalent, and no source for the information required to distinguish them existed.

Similar processes continue to produce information as the universe develops. The classic example is Bénard cell convection in a fluid gently heated from below. If other parameters making up the Rayleigh number are kept constant, the onset of convection depends on the temperature gradient through the fluid. The system is too difficult to analyse at the microscopic level, so standard treatments compare the dynamics of conduction with those of convection, and determine the transition Rayleigh number from the convergence of the equations at the onset of instability[5]. The condition of conduction is symmetrical at all levels (except for statistical fluctuations), but the onset of convection introduces new symmetries at the macroscopic level, while at the same time destroying some of the microscopic symmetry (some microscopic automorphisms are ruled out by the macroscopic movement of molecules in the fluid).

Sympatric (same locale) speciation can occur via similar processes, as can tissue differentiation during biological development (Collier et al, in press). In both cases differentiation, a form of symmetry breaking, is not under the full control of either exogenous or endogenous information. Chance fluctuations play a central role in the exact pattern of information formation. There is increasing evidence that much biological information is produced by such symmetry breaking processes (Brooks and Wiley, 1988; Kauffman, 1993; Collier and Siegel-Causey, in press). One of the major mechanisms that has been proposed to drive this sort of process is self-organisation permitted by non-equilibrium processes, where non-equilibrium conditions are maintained by a continually expanding state space within a branch system (Brooks and Wiley, 1988). There is a very strong analogy to the formation of information through symmetry breaking in cosmological processes.

It has been proposed that at least some perceptual systems (smell and hearing) involve chaos in the transducer and associated neurons (a condition of symmetry) that is driven by a sensory signal into a particular harmonic orbit of the many quasi-harmonic orbits that compose the chaotic condition (Skarda and Freeman, 1987). If all perception works similarly, then perception is a form of symmetry breaking that produces perceptual information. The information, in this case, comes from outside the perceptual system. However, from inside the system the original condition is statistically symmetrical, and the asymmetry is produced by the internal dynamics of the system under the influence of a relatively small driving force. The response of the system, if these controversial proposals are correct, is much stronger than the stimulus producing it. Perhaps the same principles can be extended to cognition, as well as perception. At present this is pure speculation, but it is certainly possible that concept formation and theory formulation in the face of incomplete evidence is also partially driven by self-organising processes that provide information beyond that contained in the evidence available. If so, symmetry breaking is fundamental to many cognitive processes as well as to perception.

To recapitulate, the original condition of the universe is statistically uniform, and hence entirely symmetrical. This statistical uniformity implies an equilibrium state (at least locally), which further implies that the early universe did not contain any information. Information, therefore, must have arisen through contingencies. The only process we know of that can produce new information from contingencies is symmetry breaking through phase separation in a system that is out of equilibrium, thereby forming branch systems. Similar branching is repeated at smaller scales as the universe differentiates and forms new branch systems (Layzer, 1990). This process can either be driven by forces externally applied to a system far from equilibrium that drive the system to self-organise to a more stable state, or can arise endogenously through self-organisation

[5]The equations and a physically motivated discussion of Bénard convection and its implications for self-organisation appear in (Collier et al, in press). The classic analysis is by Chandrasekar (1961).

resulting from non-equilibrium conditions maintained through an expanding state space within a branch system. In both cases, the appearance of new information depends on symmetry breaking permitted by the promotion of microscopic fluctuations on a statistically uniform background under conditions that are not at equilibrium. Non-equilibrium conditions permit symmetry to be broken, by allowing the sort of contingency that Weyl requires to overcome the universal symmetry implied by natural laws alone.

## 4. WHY SYMMETRY IS INTERESTING

Given that information production is symmetry reduction, the problem of the interest of symmetries becomes acute. Not all symmetries are equally interesting. Many symmetries are purely superficial, being either convenient or accidental, traceable to equilibrium conditions (which wipe out information), or both. These symmetries have little logical depth, and are relatively easy to produce. Examples are symmetries found in crystal structures in typical minerals (not the so-called "aperiodic crystals" that are found in complex biological molecules), many utilitarian human artefacts, and the obvious symmetries, radial and bilateral, found in animals and some plants[6]. The main interest of these symmetries is in the information reduction produced by our knowledge of them, together with possible advantages of having the redundant systems that these superficial symmetries imply. This is not incompatible with our having an aesthetic appreciation, or even fascination, for superficial symmetries, but it does mean that this appreciation will never penetrate deeply.

Deeper symmetries, however, are left after symmetry breaking process produce information at higher levels. These symmetries are much more interesting. Their interest, I propose, is in what they reveal about the common source of different informative structures that arise from them. The deep symmetries highlight the individuality of derivative structures, and also make their common nature more clear. Their recognition allows us (and forces us) to see the sorts of symmetry breaking – information – that make up the different structures emanating from a common source. The symmetries behind natural laws don't merely provide a convenient shorthand to compress our empirical data; they also help to highlight both the particular and the general in that data in an especially insightful way.

To take one example, the Law of Conservation of Energy (the First Law of Thermodynamics), like all conservation laws, is a symmetry law. As a universal physical law, it requires that the possible physical transformations are restricted to the ones that maintain the symmetry implied by the law. Newton's laws themselves do not imply conservation of energy (dissipative forces are not ruled out), though each of Newton's laws implies other symmetries (consider, for example, the law of reaction). Additional assumptions invoking symmetries at least as strong as conservation of energy must be made in order to derive the energy conservation law from Newtonian mechanics. One of the simplest such assumptions is that all forces are conservative. This principle imposes a condition on systems that the sign of the time parameter in a dynamical description of a system can be substituted with its opposite, and the resulting

[6]Radial and bilateral symmetry are not themselves adaptations, but are left over from the spherically symmetrical structure of primitive organisms (phylogenetically), and of the fertilised ovum (ontogenetically). Adaptations produce deviations from this ur-symmetry.

system is equally possible physically. The condition also requires that any temporal translation of a system will make no difference to its dynamics. These requirements greatly simplify the set of possible physical systems, and they worked well with atomic and molecular systems, as well as loosely coupled systems like the solar system. The recognition that heat is a form of motion of microscopic particles under the influence of conservative forces allowed the extension of the principle to dissipative systems. The principle underlies the applicability of the elegant Hamiltonian formulation of mechanics, which has been extended to fields, waves and quantum systems. The generality of the Hamiltonian formulation also allows the introduction of generalised coordinate systems, which further extends its applicability and ease of use.

Conservation of energy is a simple principle, but it is not at all obvious how it is to be applied to all physical systems. In fact it took many years to give an adequate account of thermodynamic systems in mechanical terms, and the account still has some problems. The striking thing is that such a wide variety of systems, which appear quite different on the surface, can be described by the same basic equations, all on the assumption that the total energy, kinetic and potential, of an isolated system is a constant. At the same time as the systems are taken to obey the same basic symmetry, their surface structure can be quite different. The application of the Hamiltonian formulation highlights both the underlying common characteristics of all physical systems, but the details of the application in each case highlights their differences.

The energy conservation principle is quite deep, since it is often not apparent how to apply the principle to individual cases. Some possible physical systems are so complex that there is no way to apply the principle analytically, and the best we can do is to make approximations. Nonetheless, the unifying power and simplicity of the principle leave little doubt that it should apply to all physical systems, even if this cannot be proven rigorously. In fact, the very difficulties encountered in applying the principle to certain types of systems helps to explain the nature of those systems, and how they differ from systems to which the principle can be more easily applied.

In closing, I would like to speculate that the aesthetic value of symmetries (aside from their artistic use to point to other things) is a reflection of the basis of their scientific interest (a symmetry underlying symmetry!) Shallow symmetries are boring and repetitive; too much symmetry is bland. When surface characteristics are underlain by a recognisable deep symmetry, however, the symmetry ties together the individual parts without denying them their individual interest or identity. At the same time as it gives an overall harmony and unity, deep symmetry highlights the individuality and unique interest of the parts of a composition or natural phenomenon. Appreciation of deep symmetries allows a much richer experience than would be possible otherwise. Although evolutionary biologists have tried to explain our taste for symmetry in terms of the health of potential mates, I suspect that its basis is much more deeply embedded in our nature, and in our relation to the world at large.

## REFERENCES

Aleksandrov, A. D., Kolmogorov, A. N. and Lavrentov, M. A., eds. (1969) *Mathematics: Its Content, Methods and Meaning*. Translated from Russian by Gould, S. H., Hirsch, K. A. and Bartha, T., Cambridge MA: MIT Press.

Banaschewski, B. (1977) On G. Spencer Brown's laws of form, *Notre Dame Journal of Formal Logic*, **18**, 507-509.

Bar-Hillel, Y. (1964) *Language and Information*, especially Chapters 15-17. Reading MA: Addison-Wesley.

Brillouin, L. (1962) *Science and Information Theory*, second edition, New York: Academic Press.

Brooks, D. R. and Wiley, E. O. (1988) *Evolution as Entropy: Toward a Unified Theory of Biology* 2nd edition. Chicago: University of Chicago Press.

Chaitin, G.J. (1975) A theory of program size formally identical to information theory, *Journal of ACM*, **22**, 329-340.

Chandrasekar, S. (1961) *Hydrodynamic and Hydromagnetic Stability*, Oxford: Clarendon Press.

Collier, John D. (1990) Intrinsic information, In: Philip Hanson, ed., *Information, Language and Cognition: Vancouver Studies in Cognitive Science, Vol. 1*, Vancouver: University of British Columbia Press.

Collier, J. D., Banerjee, S. and Dyck, L. (In press) A non-equilibrium perspective linking development and evolution, In: Collier, J. and Causey, D. S., eds., *Between Order and Chaos: Studies in Non-Equilibrium Biology*.

Collier, J. D. and Causey, D. S., eds. (In press) *Between Order and Chaos: Studies in Non-Equilibrium Biology*, John Baltimore: Johns Hopkins University Press.

Kauffman, S. A. (1993) *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford: Oxford University Press.

Kolmogorov, A. N. (1965) Three approaches to the quantitative definition of information, *Problems of Information Transmission*, **1**, 1-7.

Kolmogorov, A. N. (1968) Logical basis for information theory and probability theory, *IEEE Transactions on Information Theory*, **14**, 662-664.

Layzer, D. (1990) *Cosmogenesis: The Growth of Order in the Universe*, New York: Oxford University Press.

Li, M. and Vitànyi. P. (1990) Kolmogorov complexity and its applications, *Handbook of Theoretical Computer Science*, edited by J. van Leeuwen. Dordrecht: Elsevie.

Penrose, O. (1970) *Foundations of Statistical Mechanics: A Deductive Treatment*, Oxford: Pergamon Press.

Shannon, C. E. and Weaver, W. (1949) *The Mathematical Theory of Communication*, Urbana: University of Illinois Press.

Skarda, C. A. and Freeman, W. J. (1987) How brains make chaos in order to make sense of the world, *Behavioural and Brain Science*, **10**, 161-195.

Spencer Brown, G. (1972) *Laws of Form*, New York: Bantam.

Weinberg, S. (1977) *The First Three Minutes: A Modern View of the Origin of the Universe*, New York: Basic Books.

Weyl, H. (1952) *Symmetry*, Princeton: University of Princeton Press.