

PERSPECTIVE

On the logical relationship between natural selection and self-organization

G. A. HOELZER,* E. SMITH† & J. W. PEPPER‡

*Department of Biology, University of Nevada Reno, Reno, NV, USA

†Santa Fe Institute, Santa Fe, NM, USA

‡Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA

Keywords:

adaptation;
 complexity;
 entropy;
 natural selection;
 self-organization;
 thermodynamics.

Abstract

Most evolutionary biologists cherish Darwin's theory of natural selection (NS) as *the* process of adaptive evolution more than 140 years after publication of his first book on the subject. However, in the past few decades the study of self-organization (SO) in complex dynamical systems has suggested that adaptation may occur through intrinsic reorganization without NS. In this study, we attempt to describe the logical framework that relates the general process of SO to the specific process of NS. We describe NS as a mechanism that coordinates the coevolution of species in an ecosystem to effectively capture, process and dissipate solar energy into the earth's shadow. Finally, we conclude that NS is an emergent process founded on the same thermodynamic imperatives that are thought to underlie all SO. This perspective suggests that the theory of self-organizing systems offers a broader physical context in which to understand the process of NS, rather than contesting it. It even suggests the possibility that there may be a physical basis for understanding the origin of the process of NS. Rather than being merely a fluke of nature, the origin of NS that may be driven by energy flows across gradients.

Introduction

a fully adequate theory of evolution must encompass both self-organization and selection. (Coming, 1995, p. 112)

Ever since Darwin argued so persuasively about the power of natural selection (NS) in his book *On The Origin of Species* (Darwin, 1859), biologists have been quite united in the view that selection is the only natural process capable of guiding evolutionary change in an adaptive direction. However, a somewhat different, or at least different sounding, view was established in another landmark book called *What Is Life* (Schrödinger, 1945). Schrödinger's concept of adaptation as an aspect of physics has gained much attention over the past few decades (e.g. Kauffman, 1993; Depew & Weber, 1995) without having much effect on the devotion to traditional Darwinian principles by most evolutionary biologists. The emergence of complex, functional structures in

physical systems, defined with respect to the transformations a structure makes possible in its environment, has come to be called self-organization (SO; Yates *et al.*, 1987), and the purpose of this study was to explore the logical relationship between the processes of NS and SO.

Darwin took a Newtonian approach (Ruse, 1982) to explain the process of biological adaptation in which he reasoned about how the kinds of mechanical interactions occurring among individuals could cause adaptation of the population over a long time frame relative to the life span of an individual. He defined the process of NS as the tendency of some variants in a population to survive and reproduce more reliably (i.e. have greater fitness) than others, and argued that heritability of fitness differences would result in adaptive evolution. It was not until the merging of Darwinism with genetics in the late 1920s and 1930s that the important role of genetic mutation in the generation of heritable variation among individuals was widely recognized.

The current genomics revolution has revealed that our understanding of the genetic encoding of phenotypic traits was less complete than we thought. It was hard enough to imagine how *c.* 100 000 genes could

Correspondence: G. A. Hoelzer, Department of Biology, University of Nevada Reno, Reno, NV 89557, USA.
 Tel.: 775 784 4860; fax: 775 784 1302;
 e-mail: hoelzer@unr.edu

encapsulate the instructions for a human, but the sequencing of the human genome has shown that there are actually only *c.* 30 000 genes available to inform human development (Venter *et al.*, 2001). However, organismal development appears to take advantage of the information generated through SO (Camazine *et al.*, 2001) by encoding factors that create conditions favouring particular pathways of SO. If there is a fitness cost to having larger genomes (e.g. see Yedid & Bell, 2001), then it is reasonable to expect that NS would favour taking advantage of the opportunity for economical encoding made possible by the tendency of physical systems to self-organize (Camazine *et al.*, 2001; Meinhardt, 2001; Rocha, 2001). The alternative would be to eschew the fitness advantages of morphological SO and pay the fitness costs associated with actively preventing genetically non-encoded tendencies to organize as energy flows through the developing organism.

The processes of reproduction and death within a population of entities (individual organisms in the classical Darwinian view) are central to the traditional notion of NS. Together they create the mechanism for the replacement of less-well adapted forms by better adapted ones. Consequently, NS confers on populations a propensity for increasing mean fitness, which will be referred to as fitness maximization in this paper. NS helps us to understand the persistence of life because the maximization of fitness based on selective filtering of heritable variation in the past results in generally well-adapted forms in the near future. The temporal correlation of relative fitnesses results in a natural teleology in the sense that organismal processes affecting fitness, such as development, energy consumption, mating in sexual organisms and so on, are designed to function for the sake of fitness. To biologists the principle of fitness maximization appeared to distinguish living systems from nonliving ones; the latter have typically been viewed by biologists as nonadaptive.

Note that the sense in which the term maximization is used here does not necessarily lead to a maximum value, which would have no sensible meaning in the context of fitness, but it is distinct from an optimizing process, which under some circumstances could cause the parameter value of interest to decrease in order to achieve an optimal balance among interacting forces. This sense of maximization always augments the focal parameter relative to the value it would otherwise have obtained. There can be, however, a physical relationship between a maximizing process and optimizing processes, because maximization of one parameter can require coordinated changes among interacting phenomena that influence the value of that parameter. For instance, epistasis may require compensatory changes among several traits before the potential fitness benefit made possible by a change in a single trait can be fully realized. In this way maximizing processes can naturally generate complex optimizing processes (e.g. see Helbing & Vicsek, 1999).

We refer to this as an optimizing imperative because it is caused by the operation of a maximizing process. Indeed, all optimizing processes are instantiated by maximizing processes from the perspective of this conceptual framework.

Curiously, a seemingly unrelated maximizing principle was developed in the field of physics at about the same time the theory of NS was introduced called the second law of thermodynamics (Clausius, 1850). Physicists have expressed this physical law in many forms over the years, but the most common conceptualization of it, at least among biologists, is that entropy (disorder) can only increase in closed systems, including the universe as a whole (Brooks & Wiley, 1988). This law has served as the basis for understanding how some open systems, specifically complex dynamical systems, can self-organize (Schrödinger, 1945; Nicolis & Prigogine, 1989). One way to recognize SO systems is that they can evolve from a disorganized state to a more organized state due to internal system dynamics without the imposition of pattern from external forces (Camazine *et al.*, 2001). Systems that originate in response to, and are maintained by, the optimizing imperative of the second law of thermodynamics are sometimes called dissipative structures (Nicolis & Prigogine, 1989) or complex adaptive systems (Levin, 1995). In contrast to equilibrium thermodynamics, which identifies the stable macrostates of nature as those that maximize certain configurational entropies, subject to whatever constraints are applied to the system (Jaynes, 1983; Gell-Mann & Lloyd, 1996), the proposal for dissipative structures is that dynamically stable states are those that maximize the *rate* of entropy increase in the universe relative to accessible alternative configurations (Ozawa *et al.*, 2001). Considering the possibility of a universal rate maximizing law introduces notions of function that may serve as the basis for processes of SO. Indeed, it has been shown that SO systems can exhibit autonomous structural adaptation to changes in environmental conditions preserving their function in this sense (Helbing & Vicsek, 1999).

The similarities between physical SO and biological evolution under NS described above are striking. Indeed, it is clear that the process of SO represents a potential explanation for adaptive biological evolution. Many authors writing about SO systems have made this point, most notably Stuart Kauffman (1993), but most evolutionary biologists have not accepted this proposal. We will consider possible explanations for this resistance and attempt to articulate the logical relationship between NS and SO with the aim of facilitating the integration of SO into evolutionary theory.

Why the reticence?

We think that one factor interfering with cross-disciplinary scientific communication and open-mindedness is a culture of defensive argumentation that has been

culturally imbued into evolutionary biologists in response to continual socio-political attacks on the validity of evolution science. Given the central role of Darwinism in evolutionary biology, the defensive stance has largely been a defence of Darwinian principles against a series of alternative explanations for adaptedness based on (generally biblical) beliefs in metaphysical forces (Ruse, 1982). These alternatives have often been cleverly disguised as scientific (physical and testable) alternatives, the most recent form being labelled 'intelligent design theory' (Behe, 1996). It has essentially become a matter of social responsibility for evolutionary biologists to join the battle in defence of Darwinism, but there is a scientific cost associated with this cultural norm. Alternative ways of describing evolutionary processes, complementary to NS, can elicit the same defensive posture without critical analysis. The tendency to dismiss potential alternatives has been facilitated in this case by the following two factors: (1) the theory of SO systems has been developed primarily in the fields of physics (e.g. Bak, 1996) and computational science (e.g. Wolfram, 2002), rather than from within evolutionary biology and (2) emergent phenomena that occur when local interactions drive a global system (Holland, 1999) can appear mystical to those trained in the tradition of reductionism (Wolfram, 2002).

Despite the conservative nature of evolutionary biologists in this matter, we believe most would be interested in additional ways of understanding coevolution and selection, and possibly predicting its aggregate effects from a whole-system perspective, if such were available. A review of the SO literature indicates that the discussions about the logical relationship between SO and NS have generally been vague, and that consensus has not been achieved on this issue (see Depew & Weber, 1995 for another review of the diversity of viewpoints). One commonly stated view is that NS is subordinate to SO.

Evolution ... is emergent order honored and honed by natural selection. (Kauffman, 1993)
 Natural selection does not have a lot to do except act as a coarse filter that rejects the utter failures. (Goodwin, 1994)
 When proper account is taken of the non-randomness in matter and the self-organizing properties of process, selection becomes redundant and misleading as an explanation ... (Ho & Saundes, 1986)

Others have emphasized NS as the dominant process.

it is obvious that thermodynamic considerations must be one of the priorities of natural selection.
 Living systems must ... submit to the ultimate 'censorship' of natural selection. (Comins, 1995, p. 95)
 In biological systems, self-organization reveals itself as a powerful mechanism used by natural selection for the creation of diverse regulatory and morphogenetic processes. (S. Camazine: <http://www.scottcamazine.com/personal/research/index.htm>)

In comparing these comments, it is important to distinguish the scale on which each is predicated. Widely recognized opportunities for SO in ontogeny are created by the genetically programmed introduction of particular cell types and interactions (while of course an infinity of other conceivable opportunities for SO are ruled out by the nonexistence of appropriate cell types and interactions). In this narrow context SO is a component of development subordinate to NS, which acts in part on the ability of the organism to exploit SO for efficiency.

Subordinate SO in development does not rule out super-ordinate SO at the ecological level, whose projection onto the dynamics of individual organisms we describe as NS. Indeed, the distinction between subordination and super-ordination itself may be artificial, as it is not clear which instances of SO in development (or allometry, as we discuss below) may have superseded individual dynamics, and served as a framework around which the residual mechanisms of change by NS collected.

Carlson and Doyle in their work on highly optimized tolerance (e.g. Carlson & Doyle, 2000) suggest that NS (and other sources of engineering) results in structural designs that constrain or guide the process of SO, whereas others have suggested that SO is constrained by NS. For example, consider the following quote:

natural selection ... is not free to evolve any organism, but it is constrained by the self-organizing properties of the materiality of the organisms it acts upon. (Rocha, 1998, p. 346)

Finally, some authors have implied that NS and SO are not as distinct from one another as they might appear.

We have documented that ecological processes are driven and governed by thermodynamic imperatives ... We believe that these same principles extend to Darwinian selection. (Schneider & Kay, 1994, p. 45)
 models based on phenotypes, and including nonlinear and collective effects, suggest that evolution can indeed be viewed as a process whereby the ecosystem self-organizes. (Stewart, 2003, p. 1101)

Despite the apparent contradictions among these quotes, we think that they may all be valid statements representing aspects of the complex relationship between these processes. A more comprehensive analysis of the relationship between SO and NS is needed to reveal how each of these viewpoints fits into the bigger picture.

On the physics of SO systems and its application to life on earth

The understanding and modelling of SO in physics are still in early stages. Although it is widely appreciated that equilibrium statistical mechanics is inadequate to account for dynamic pattern formation, SO can occur

in systems whose instantaneous configuration may be either very near thermal equilibrium, or manifestly very far from it. There are thus at least two views of the physical basis of SO. Most of the research on SO systems has focused on conditions far from equilibrium (e.g. see Nicolis & Prigogine, 1989; Ozawa *et al.*, 2001), but in cases where the equilibrium entropy remains a well-defined and usable concept, SO of a subsystem has been shown to increase the rate of entropy transport, though no net entropy is produced (Malkus, 1956; Smith, 1998, 1999). Although an integrated theory of the various instances of SO has yet to be built, commonalities among the models and empirical examples that have been explored (Ozawa *et al.*, 2001) indicate that the potential for emergence of SO dynamical structures exists wherever there are sufficiently concentrated spatial gradients in free energy. This potential is called a *voltage* when dealing with electrical energy, but the same concept applies to free energy in any form. The function of the SO system is to channel the flow of energy across the gradient, thus reducing the steepness of the gradient. The structure and function of the SO system is both created and sustained by extracting work from the free energy differences as energy flows through the channel provided by the SO system. Thus, for example, the plasma channel of a lightning strike is created by the pressure to relieve dielectric stress on the atmosphere, formed through charge transport in convective storms.

Other familiar examples of physical SO systems that evidently work this way include the convection cells themselves (within storms) and vortices (tornadoes) forming in the atmosphere when heat builds up at the earth's surface. When the temperature gradient between the lower and upper atmosphere becomes sufficiently great, simple conduction of heat through the atmosphere is no longer rapid enough to maintain structural stability of the air mass and convection cells form, which channel the flow of heat and increase the rate of heat exchange across the gradient (Ozawa *et al.*, 2001). If the gradient becomes so severe that the effect of convection cells is not adequate, then there is a potential for forming the yet more effective flow structure of the vortex. It seems easy to identify where the gradient exists in the formation of weather, but it is not as obvious that the vertical temperature gradients found in the atmosphere would not exist to nearly such an extent if the earth did not spin relative to the position of the sun. Heat builds up at the surface due to solar radiation on the sunny side much faster than it builds up in the upper atmosphere. The heat energy is stored by the thermal mass of the surface and conducted into the atmosphere at a rate dependent upon the steepness of the thermal gradient, which increases as the surface is transferred to the dark side of the planet. In addition, heat is passed into space from the upper atmosphere on the dark side of the earth much faster than it is conducted into the upper atmosphere from the surface, which sustains the gradient in the atmosphere

that fuels weather systems into the night. Consequently, the amount of weather would be greatly reduced, and probably concentrated in the band between the light and the dark, if the earth did not spin relative to the sun. The importance of planetary spin reveals a more extreme, larger scale energy gradient that is also diminished by weather: the spatial gradient caused by the earth's shadow. Solar energy is relatively dense in the space between the sun and the earth, whereas there is an energy vacuum in the earth's shadow. The accumulation of energy on the sunny surface of the earth further intensifies the intensity of this localized gradient, which would not exist at all if the earth was not present as an obstruction to the flow of radiant energy. Weather channels some of this energy across the surface, making the earth more transparent than an equivalent weatherless planet to solar radiation. Note that the main message of this section is not analogy, but the embedding of NS as one ingredient of SO at Earth scale.

Solar radiation is also spectrally out of equilibrium with the background microwave radiation of space. Not only earth, but even the vacuum, create a 'shadow' of a different kind, to the extent that they provide poor (or no) mechanisms for scattering of energy from visible to microwave bands. The process of photosynthesis, the primary energetic input to earth's biomass today, creates a conduit across the spectrum by absorbing visible light and converting it to chemical energy, which ultimately re-emerges in heat through processes of catabolism and decay, and through so-called 'pigment effects', a generic name given to microwave fluorescences that may be a necessary aspect of irreversible absorption in photosystems. Most of the free energy change from this transduction is obtained in going from the effectively 6000 K black body of sunlight to the earth's 300 K microwave background. (A small residual potential, from transducing earth's 300 K background to the 3 K background of space, is unavailable to life because the atmosphere is roughly in radiative equilibrium.)

Wald (1974) has emphasized that the problem of repeatable absorption of sunlight is chemically challenging and appears to have been solved only twice by life, with the synthesis of the chlorophylls and the rhodopsins. Studies of horizontal transfer of photosynthesis-related genes (Raymond *et al.*, 2004) further suggest that refinements in pigments were rare and valuable enough to be widely exchanged among taxa, rather than independently discovered. In contrast to the complex apparatus of photosystems, atmospheric scatterers tend to be small molecules, unsuited to cross-band fluorescence because of quantum selection rules. Dry land scatterers are also a modest contribution to the earth's albedo, suggesting that opportune fluorescent transitions have largely been removed by photodissociation. (We will return in a moment to the main source of albedo, the oceans, whose unusual physical chemistry is connected to weather but not energetically to life.)

In the Appendix we show, from simple considerations of energy and momentum conservation, that (inelastic) light absorption transfers roughly one billion times as much energy per event from the visible to microwave bands, as competing elastic scattering processes, when the atoms participating in the two processes cycle between comparable (low) temperatures. An aggregate consequence of photosynthesis is therefore that the effective specific heat, the energy absorbed per degree of warming, of the earth as a whole is increased. Specific heat is the property that makes a spinning earth into a heat engine, transferring energy not only across spectral bands but also from the sunny to the dark side.

Most abiotic conversion of visible to thermal energy on earth occurs through diffuse, multiple scattering of light in the oceans, made possible by the unique physical properties of water. This is the major source of energy driving global-scale weather and climate, but it is incapable of generating the structures of life because of the small energy capture per molecule and per scattering event. Photosynthesis, while smaller in net energy capture because of the tiny mass of life compared with the mass of the oceans, contributes a parallel channel by preserving large fractions of the energy from visible photons in individual molecular bonds (Blankenship, 2001). The energy and entropy flows through this parallel channel, inaccessible to weather, provide the source of free energy to create and maintain biochemical networks capable of supporting photosynthesis.

The transduction of sunlight not only makes photosynthetic life possible, but also it appears to have stabilized its core chemistry even in the face of major shocks to planetary ecosystems, such as those resulting in recurring mass extinctions. This universally known but underemphasized fact suggests that photosynthesizing life is a statistically favoured component of the biosphere, or that a high-flux channel for light transduction is a favoured endpoint, towards which perturbed ecosystems recover. Physical principles expressed in terms of stable end states imply a natural teleology, which we have suggested (somewhat imprecisely) is the reduction of the spectral and spatial energy gradient imposed by the situation of earth in a solar radiation bath.

However, it is unlikely that life could have served this function at its inception, because the earliest organisms apparently did not photosynthesize (Blankenship, 2001). An alternative energy gradient also exists due to geochemical cycling driven by fission in the earth's interior, which exchanges heat and gravitational energy for chemical energy. Reduced states of heavy metals are convected downward in the heat transfer cycle, and are pyrolytically re-oxidized by seawater when they are carried back up at sites like hydrothermal vents, creating reductants and other energetic chemical species like phosphate esters and thioesters. Geochemical cycling was more intense on the early earth than it is today, and the absence of biologically produced molecular oxygen per-

mitted more favourable conditions for local formation of reducing environments. Numerous ecosystems have been found at such locations today, which harvest the reliable energy-rich molecules and extract energy as well as anabolic precursors from the redox potentials and functional groups in these compounds. It is not established, though it seems likely, that geochemical cycling even today provides a sufficient potential for the long-term maintenance of living systems. It remains a possibility, though, that all such extant systems are ephemeral overflows from ecosystems sustained by solar energy; recent data do show that most species found in deep sea hydrothermal vents and cold seeps are relatively recently derived from solar-energy-based systems (Van Dover *et al.*, 2002).

Do NS and SO really optimize different criteria?

The definitions of the terms NS and SO suggest that a primary distinction between these concepts, if not the actual processes, is that different criteria are maximized; NS maximizes fitness whereas SO maximizes (for example in certain ecologies) the rate of energy flow across a gradient. However, these criteria overlap far more than is superficially apparent. For example, both fitness and the channelling of energy flow require consumption of energy and the extraction of work to create structure and function. It is entirely consistent with both theory and observations in ecology and evolution that NS would generally result in more effective energy processing at any level of biological organization (the cell, multicellular organism, population, ecosystem), so it seems plausible that the maximization criteria of SO and NS are at least highly complementary. The question then becomes 'is it possible for NS and SO to conflict with one another?'

There are reasons to think that NS can occasionally drive populations in directions that are inconsistent with long-term thermodynamic 'goals'. For example, some models suggest that sexual selection can sometimes reduce population sizes and lead to eventual extinction of the species (e.g. Houle & Kondrashov, 2002). This would certainly be an example of NS acting in a way that hinders persistence of energy processing in a particular system. However, there are two points that are important to consider in such cases. First, individuals in such systems are likely to become ever more effective energy processing machines as the population goes to extinction. They would bear energetic costs of producing more flamboyant epigamic traits or morphological weaponry, intense courtship activity and/or intrasexual combat. SO systems need not behave perfectly with regard to their future utility as energy processors; we would not describe the processes that build hurricanes as flawed just because hurricanes occasionally wander over cold water or land, which result in the demise of those processes. The

continual proliferation of storms, and biological lineages, sustains these processes at the global scale. This leads to the second point; NS would also act in a compensatory fashion within the ecosystem by favouring acquisition and processing of the free energy no longer channelled into the diminished or extinct population. If so, NS could promote effective energy processing at the global scale. Ultimately, NS (at the level of the individual) is the projection onto the individual, of a physical process of differential growth under competitive exclusion, modulated by the constraints of the individual as a unit of selection and the population as a reservoir for inheritance. The rate of the underlying growth is proportional to the strength of the energy gradient multiplied by both the rate of flow through it and the efficiency of the phenotype in its constructive use of that energy.

The origin of life and its effects on thermodynamic processes: SO and the origin of NS

As with all emergent dissipative structures, organisms become 'atoms' of interaction, whereas smaller entities (e.g. macromolecular polymers such as RNA, DNA and proteins) previously characterized the maximum sizes of particles interacting in the setting of prebiotic biochemical evolution. The emergence of higher scale functional structures in SO systems makes thermodynamic processes more coarse-grained in space.

Another quality characterizing all living systems on earth, which is also a fundamental basis for the process of NS, is the birth/reproduction/death cycle. Salthe (1993) has suggested that nonliving SO systems, like atmospheric convection cells and vortices, also tend to exhibit birth, senescence and death, but reproduction is a less critical aspect of these systems in which birth through spontaneous generation is more common; therefore, persistence of the phenomenon of weather is not as dependent on reproduction as the phenomenon of life. One could say the same thing from another point of view. The sophistication of cellular life far exceeds any plausible sophistication of the first biochemistry on the road to life. Thus we would never see life forms emerging *de novo* surviving in competition with modern forms, and by far the favoured route to produce more life is by reproduction of existing life. In comparison, the ability of weather structures to increase in complexity and efficiency is very limited; indeed it is difficult to imagine any other substrate than chemistry capable of supporting as broad a range of complexity as likely separated primordial from modern life. Thus the dominance of a mature hurricane is sufficient to exclude the spontaneous generation of others in the immediate vicinity, but not to be the sole source for them in all places at all times (though note that the wall replacement cycle has elements of reproduction of like kinds). The transient nature of biological individuals, and their continual replacement in

the population through reproduction, yields temporal quanta of interaction and a temporal grain to system dynamics.

In general, the coarseness of interacting grains in thermodynamically driven SO systems is an important factor because coarser grains mean smaller interacting sets and increased stochasticity, which is the source of creativity underlying adaptive structural changes. We are familiar with the role of mutation as a source of random variation needed to fuel NS, and there is an optimal rate of mutation that favours adaptive evolution under NS (Iwasaki & Yonezawa, 1999), just as there is an optimal level of stochastic perturbation favouring adaptive responses of thermodynamic SO systems (Helbing & Vicsek, 1999). There is abundant evidence that NS is effective at regulating both body size and generation time (Stearns, 1992), so NS could be a useful mechanism for tuning the degree of stochasticity in systems to satisfy thermodynamic imperatives.

Natural selection shares another important feature in common with the process of SO; it persists due to the positive feedback that is inherent to the process. Because NS results in adaptation, the dynamical system that gives rise to the process of NS endures. The process of NS feeds on its own consequences, the production of more replicators, which is why it is easier to understand NS as a fundamental process of evolution than it is to understand why the conditions that gave rise to NS came about. It is our view that the latter issue can be appreciated by considering the way process emergence generally occurs in complex dynamical systems.

Adaptation of a population vs. adaptability of an entity: NS as a mechanism of SO

Evolutionary biologists have traditionally distinguished between the developmental processes of adaptive phenotypic plasticity, where an individual adjusts its structure to maintain optimal functional performance, and adaptation, which requires mutation and selection within a population of reproducing individuals. Indeed, adaptive phenotypic plasticity has been treated almost exclusively as the result of organismal design by NS (e.g. Via, 1993; but see Emlen *et al.*, 1998). Nevertheless, adaptability has been understood as an autonomous process of adaptive change that does not involve mutation and selection among reproducing agents in a population. A similar process is also widely recognized to occur at the ecosystem level. Each species plays a particular role in a functioning ecosystem, and evolution is constrained by NS to maintain the coordination of the system as a whole. For example, the study of food web architectures shows that NS in coevolutionary setting prevents the evolution of communities composed entirely of predatory species, and a flourishing community of primary producers inevitably attracts or generates grazers, then

predators and so on (Pimm, 1982). Thus, NS is a mechanism that coordinates the coevolution of populations so as to promote homeostasis and efficient energy processing at the ecosystem level. The dynamical maintenance of ecosystem structure allows ecosystems themselves to adjust adaptively to changes in their environments, such as would occur through global climate change. This could be described as the adaptability of an ecosystem (see Salthe, 1985), because it does not involve competition among ecosystems. Ultimately, NS in the context of coevolutionary systems serves as a mechanism that coordinates the internal dynamics of ecosystems in a way that maximizes the capture and processing of solar energy.

Prospects for an empirical research programme

Broadening our conceptual framework is a worthwhile endeavour, but the real value of such an exercise can only be realized through empirical evidence and application; therefore, we strongly recommend that researchers collect data addressing questions that arise out of this new perspective. We think that the potential value of such an approach is illustrated by a recent study in which Moses & Brown (2003) reveal the scaling relationship between number of offspring born per female (fertility), a reasonable proxy for fitness, and the rate of energy consumption (metabolic rate not adjusted for body size). The theory on which this study is based, metabolic scaling theory (Brown *et al.*, 2000, 2004), is founded on the notion that structures of systems guiding the flow of materials that constitute metabolisms are organized into fractal branching networks because this is the most efficient geometry for material delivery to or removal from a space. The optimization of distribution networks that underlies metabolic scaling theory may arise through abiotic SO, as appears to be the case with the fractal branching observed in the structures of lightning and river drainages. Alternatively, it may arise biotically through the process of NS. As predicted by metabolic scaling theory, reproductive rate (R) scales as metabolic rate (B) to the power of $-1/3$ ($R \propto B^{-1/3}$; West *et al.*, 1999); thus, organisms that process energy at greater rates produce fewer offspring. Some might find it surprising from a strictly Darwinian point of view that the more energy an organism obtains the fewer offspring it produces, as a general rule, and that the slope of this relationship is well predicted by a model that is not based on the theory of NS. Even more surprising is the fact that human cultures fit the relationship that holds for all other mammals only if extra-physiological energy consumption (e.g. for manufacturing and running appliances) is included. Our reproduction is seemingly controlled to a large degree by our individual capacities for energy processing. This relationship must be central to an understanding of fitness, yet the theory of NS has

not led either to an examination of this relationship or to a quantitative prediction of the slope. Eventually, we believe that these sorts of data will make the case for a broader conceptual framework of NS most convincing. For now, we argue that the study by Moses & Brown (2003) both illustrates the kind of research question that is motivated by the perspective we advocate in this study and provides some empirical support for its validity.

NS and SO – a synthesis

Ecologists and evolutionary biologists often distinguish between proximate and ultimate causes as if there were only two valid levels of explanation. For example, the bird flew to escape a terrestrial predator (proximate), or because NS caused the evolution of a flying morphology in its ancestors that has retained its adaptive value in this species (including that ability to escape terrestrial predators; ultimate). We suggest that there is a broader spectrum of levels at which particular observations can be explained that goes beyond the reaches of traditional proximate and ultimate explanations. We offer the conjecture that the truly ultimate explanation for any dynamical event, and the qualities of any emergent dissipative structure, including organisms and ecosystems, is thermodynamic. The so-called ultimate explanations based on the emergent process of NS are more fundamental than the so-called proximate explanations, but they are not the most fundamental ones (see Salthe, 1993). Similarly, most proximate explanations in biology are usually far from the most reduced cause and effect mechanism that could be considered (e.g. the mechanical forces among interacting molecules). A schematic representation of this extended spectrum of proximate and ultimate explanations is illustrated in Fig. 1.

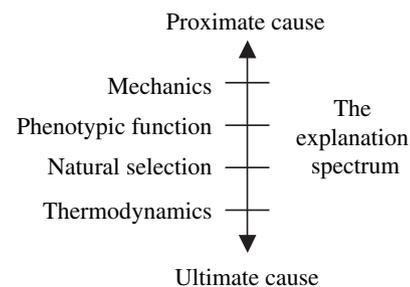


Fig. 1 A schematic view of the spectrum of explanatory levels from ultimate to proximate. The levels commonly called ultimate and proximate by ecologists are shown in the middle of this list as natural selection (NS) and phenotypic function respectively. Purely mechanical explanations would typically be more proximate in nature than explanations based on phenotypic function, and thermodynamic explanations are more fundamental than those based on NS.

This paper is a contribution towards a more comprehensive understanding of the relationship between the processes of NS and SO. Despite seemingly important distinctions relating to optimization criteria and mechanisms, close inspection reveals that these processes are largely, if not entirely, consistent and complimentary with one another. Our informal analysis specifically suggests that NS is an emergent process founded on the thermodynamically driven origin of a birth/reproduction/death cycle (see also Holland, 1976). At a grander scale, NS acting within the coevolutionary context of ecosystem dynamics is an effective mechanism for optimizing the channelling of solar energy across the spatial gradient caused by the inherent opacity of the earth. The notion of an emergent process may seem unfamiliar, so a simple analogy might help to summarize and synthesize many of the points made above: NS is to SO as language is to communication. Language is not synonymous with communication, yet it is a quantum leap forward in communicative effectiveness. It is also not merely a more complicated form of simpler modes of communication. It is a mechanism of communication that has given rise to a diversity of details (form and function) not possible in the absence of language. In sum, language is a manifestation of communication, yet it is also something more than communication can be in the absence of language. Language has emerged from the drive to communicate just as NS has emerged from the universal drive to break down gradients. For a thorough discussion of this view of communication see Deacon (1997).

The framework materializing from these considerations indicates that the physics of SO systems provides a more general context in which to understand NS. It could help us understand the origin of living systems and the qualities that made NS an inherent property of those systems. We submit the conjecture that NS is no more and no less than any other thermodynamic process and that all emergent processes (e.g. convection) are the equivalent of cryptic (some more cryptic to our eyes than others) thermodynamic games resulting in the higher order processing of fuel (gradients). This perspective aims to place NS into the bigger picture of thermodynamic evolution of the universe at large (Brooks & Wiley, 1988; Schneider & Kay, 1994).

Acknowledgments

We would like to thank H. Bagheri-Chaichian and J. Brantingham for helpful discussions during the development of these ideas. We would also like to thank H. Bagheri-Chaichian, Y. Chaver, T. Deacon, J. Lyons-Weiler, J. McGinn, W. Morse, T. Nickles and S. Salthe for comments on earlier drafts of the paper. Discussions that led to this paper occurred while the all authors were working at, and being supported by, the Santa Fe Institute.

References

- Bak, P. 1996. *How Nature Works: The Science of Self-organized Criticality*. Springer-Verlag, New York.
- Behe, M. 1996. *Darwin's Black Box: The Biochemical Challenge to Evolution*. The Free Press, New York.
- Blankenship, R.E. 2001. Molecular evidence for the evolution of photosynthesis. *Trends Plant Sci.* **6**: 4–6.
- Brooks, D.R. & Wiley, E.O. 1988. *Evolution as Entropy: Toward a Unified Theory of Biology*. University of Chicago Press, Chicago.
- Brown, J.H., West, G.B. & Enquist, B.J. 2000. Scaling in biology: patterns and processes, causes and consequences. In: *Scaling in Biology* (J. H. Brown & G. B. West, eds), pp. 1–24. Oxford University Press, Oxford.
- Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. & West, G.B. 2004. Toward a metabolic theory of ecology. *Ecology* **85**: 1771–1789.
- Camazine, S., Deneubourg, J.-L., Franks, N.R., Sneyd, J., Theraulaz, G. & Bonabeau, E. 2001. *Self-organization in Biological Systems*. Princeton University Press, Princeton.
- Carlson, J.M. & Doyle, J. 2000. Highly optimized tolerance: robustness and design in complex systems. *Phys. Rev. Lett.* **84**: 2529–2532.
- Clausius, R. 1850. Ueber die bewegende Kraft der Wärme und die Gesetze die sich daraus für die Wärmelehre selbst ableiten lassen. *Annalen Phys. Chem.* **155**: 368–397: 500–524 [English translation in Mendoza, E. 1977. Reflections on the Motive Power of fire by Sadi Carnot and Other Papers on the Second Law of Thermodynamics by E. Clapeyron and R. Clausius, pp. 73–74, 109–152. Peter Smith, Gloucester.].
- Corning, P.A. 1995. Synergy and self-organization in the evolution of complex systems. *Syst. Res.* **12**: 89–121.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.
- Deacon, T. 1997. *The Symbolic Species*. The Penguin Press, London.
- Depew, D.J. & Weber, B.H. 1995. *Darwinism Evolving*. MIT Press, Cambridge.
- Emlen, J.M., Freeman, D.C., Mills, A. & Graham, J.H. 1998. How organisms do the right thing: the attractor hypothesis. *Chaos* **8**: 717–726.
- Gell-Mann, M. & Lloyd, S. 1996. Information measures, effective complexity, and total information. *Complexity* **2**: 44–52.
- Goodwin, B. 1994. *How the Leopard Changed its Spots: The Evolution of Complexity*. Princeton University Press, Princeton.
- Helbing, D. & Vicsek, T. 1999. Optimal self-organization. *New J. Phys.* **1**: 13.1–13.17.
- Ho, M.W. & Saundes, P.T. 1986. Evolution: Natural selection or self-organization. In: *Disequilibrium and Self-organization* (C. W. Kilmister, ed.), pp. 231–242. Reidel Publishing Company, Dordrecht.
- Holland, J.H. 1976. Studies of the spontaneous emergence of self-replicating systems using cellular automata and formal grammars. In: *Automata, Languages, Development* (A. Lindenmayer & G. Rozenberg, eds), pp. 386–404. North-Holland, New York.
- Holland, J.H. 1999. *Emergence: From Chaos to Order*. Perseus Publishing, New York.
- Houle, D. & Kondrashov, A.S. 2002. Coevolution of costly mate choice and condition-dependent display of good genes. *Proc. R. Soc. Lond. B* **269**: 97–104.

- Iwasaki, Y. & Yonezawa, F. 1999. A simple model of asymmetric quasi-species: network sustained by emergent high mutation rates. *Physica D* **126**: 160–170.
- Jaynes, E.T. 1983. *Papers on Probability, Statistics and Statistical Physics*. D. Reidel Publishing Co., Dordrecht.
- Kauffman, S.A. 1993. *The Origins of Order: Self-organization and Selection in Evolution*. Oxford University Press, Oxford.
- Levin, S.A. 1995. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems* **1**: 431–436.
- Malkus, W.V.R. 1956. Outline of a theory of turbulent shear flow. *J. Fluid Mech.* **1**: 521–539.
- Meinhardt, H. 2001. Organizer and axes formation as a self-organizing process. *Inter. J. Develop. Biol.* **45**: 177–188.
- Moses, M. & Brown, J.H. 2003. Allometry of human fertility and energy use. *Ecol. Lett.* **6**: 295–300.
- Nicolis, G. & Prigogine, I. 1989. *Exploring Complexity: An Introduction*. W. H. Freeman and Company, New York.
- Ozawa, H., Shimokawa, S. & Sakuma, H. 2001. Thermodynamics of fluid turbulence: A unified approach to the maximum transport properties. *Phys. Rev.* **64**: August 2001.
- Pimm, S.L. 1982. *Food Webs*. Chapman and Hall, London.
- Raymond, J., Siefert, J.L., Staples, C.R. & Blankenship, R.E. 2004. The natural history of nitrogen fixation. *Mol. Biol. Evol.* **21**: 541–554.
- Rocha, L.M. 1998. Selected self-organization and the semiotics of evolutionary systems. In: *Evolutionary Systems: Biological and Epistemological Perspectives on Selection and Self-organization* (S. Salthe, G. Van de Vijver & M. Delplos, eds), pp. 341–358. Kluwer Academic Publishers, Dordrecht.
- Rocha, L.M. 2001. Evolution with material symbol systems. *Biosystems* **60**: 95–121.
- Ruse, M. 1982. *Darwinism Defended: A Guide to the Evolution Controversies*. Addison Wesley Publishing Company, Reading.
- Salthe, S.N. 1985. *Evolving Hierarchical Systems*. Columbia University Press, New York.
- Salthe, S.N. 1993. *Development and Evolution: Complexity and Change in Biology*. MIT Press, Cambridge.
- Schneider, E.D. & Kay, J.J. 1994. Life as a manifestation of the second law of thermodynamics. *Math. Comp. Modeling* **19**: 25–48.
- Schrödinger, E. 1945. *What is Life? The Physical Aspect of the Living Cell and Mind and Matter*. Cambridge University Press, Cambridge.
- Smith, E. 1998. Carnot's theorem as Noether's theorem for thermoacoustic engines. *Phys. Rev. E* **58**: 2818–2832.
- Smith, E. 1999. Statistical mechanics of self-driven Carnot cycles. *Phys. Rev. E* **60**: 3633–3645.
- Stearns, S.C. 1992. *The Evolution of Life Histories*. Oxford University Press, Oxford.
- Stewart, I. 2003. Self-organization in evolution: a mathematical perspective. *Phil. Trans. Soc. Lond. A* **361**: 1101–1123.
- Van Dover, C.L., German, C.R., Speer, K.G., Parson, L.M. & Vrijenhoek, R.C. 2002. Evolution and biogeography of deep-sea vent and seep invertebrates. *Science* **295**: 1253–1257.
- Venter, J.C. et al. 2001. The sequence of the human genome. *Science* **291**: 1304–1351.
- Via, S. 1993. Adaptive phenotypic plasticity: target of by-product of selection in a variable environment? *Amer. Natur.* **142**: 352–365.
- Wald, G. 1974. Fitness in the universe: choices and necessities. *Orig. Life Evol. Biosph.* **5**: 7–27.
- West, G.B., Brown, J.H. & Enquist, B.J. 1999. The fourth dimension of life: Fractal geometry and allometric scaling of organisms. *Science* **284**: 1677–1679.
- Wolfram, S. 2002. *A New Kind of Science*. Wolfram Media, Inc., Champaign.
- Yates, F.E., Garfinkel, A., Walter, D.O. & Yates, G.B. 1987. *Self-organizing Systems: The Emergence of Order*. Plenum Press, New York.
- Yedid, G. & Bell, G. 2001. Microevolution in an electronic microcosm. *Amer. Natur.* **157**: 465–487.

Appendix

A simple estimate in one-dimension shows that photon capture extracts energy from visible light roughly one billion times more effectively than scattering, all of which is converted to thermal radiation by life. A visible photon can be considered to have momentum p_γ and energy $E_\gamma = p_\gamma c$, where c is the speed of light. Elastic scattering of such a photon with an atom leads to almost total recoil. The photon reverses direction and momentum essentially perfectly, and so the momentum transferred to the atom Δp_A is given by

$$\Delta p_A = \frac{2E_\gamma}{c}. \quad (1)$$

Because the atom is a nonrelativistic object, whose thermal motion has no preferred direction with respect to incident photons, its change in energy as a result of elastic scattering is just the Newtonian result. Its average (denoted by angle brackets) over the thermal motion that the atom may have in any given instance of scattering is given by

$$\langle \Delta E_A \rangle = \frac{(\Delta p_A)^2}{2m} = \frac{2E_\gamma^2}{mc^2}, \quad (2)$$

where m is the mass of the atom. Now, if such scatterings are the main source of energy of the atoms (which for abiotic matter they are), then they also set the kinetic temperature, proportional to the mean kinetic energy. One such scattering per atom in some volume therefore increases the kinetic temperature by

$$\Delta T_{A,k} = \frac{2}{3} \langle \Delta E_A \rangle = \frac{4E_\gamma^2}{3mc^2}. \quad (3)$$

As an approximation we have set Boltzmann's constant to unity and are measuring temperature in energy units allowing us to define a specific heat for this process, which in equilibrium would be the ratio dE/dT for reversible transformations. This is compatible with the equilibrium definition, but is being applied somewhat outside its domain of validity, because our definition uses explicitly time-dependent quantities. There is no functional difference in the elastic case, but there will be when we consider photon absorption, which is not a reversible transformation in the biosphere. Nevertheless, for purposes of the transport of energy across both space and the chromatic spectrum, it functions the same way. Then

$$\frac{dE_A/dt}{dT_{A,k}/dt} \leftrightarrow \frac{dE}{dT} = \frac{3}{2}. \quad (4)$$

In other words, the specific heat for this type of elastic scattering is just what it is in equilibrium thermodynamics, a factor of order unity. Suppose we ask now what happens if the atom can scatter inelastically. Rather than reverse direction, the photon simply gets absorbed, so the momentum transfer to the atom is smaller than in the elastic case by a factor of two:

$$\Delta p_A = \frac{E_\gamma}{c}. \quad (5)$$

The resulting change in mean kinetic energy is then smaller by a factor of four:

$$\langle \Delta E_{A,\text{kin}} \rangle = \frac{E_\gamma^2}{2mc^2}, \quad (6)$$

and so is the change in kinetic temperature:

$$\Delta T_{A,k} = \frac{E_\gamma^2}{3mc^2}. \quad (7)$$

Meanwhile, the expected change in the energy of the electronic state is comparable with the energy of the photon:

$$\langle E_{A,\text{elec}} \rangle = E_\gamma. \quad (8)$$

This electronic contribution far exceeds the energy gained kinetically, and is effectively the whole change

in the atom's energy. The resulting expression for what behaves as the specific heat of the matter being shuttled between light and dark is

$$\frac{dE_A/dt}{dT_{A,k}/dt} \leftrightarrow \frac{dE}{dT} = \frac{3mc^2}{2E_\gamma} \approx 10^9 - 10^{10}, \quad (9)$$

for any of the common elements. Not all of the atoms in an ecological system participate in inelastic scattering, so the actual change in the specific heat of the environmental shuttle is this billion-fold factor, multiplied by the fraction of atoms that participate in light trapping. Indeed, it is the distribution of the large energy from few events, to the heat background carried by bulk biomass, that makes this use of an 'effective specific heat' an appropriate coarse-grained physical description of the function of photosynthetic life. With such a large multiplier per event, though, even a modest fraction of terrestrial matter rendered photosynthetic by plants can produce a large enough change in energy transport to power substantial chemical machinery.

Received 31 January 2006; revised 13 April 2006; accepted 27 April 2006