Open Sys. & Information Dyn. **6**: 241–302, 1999

Printed in the Netherlands

Complexly Organised Dynamical Systems

J. D. Collier* and C. A. Hooker*

Department of Philosophy, University of Newcastle, Australia, Facsimile 049 21 6928, e-mail: plcah@cc.newcastle.edu.au

(Received: May 22, 1998)

Abstract. Both natural and engineered systems are fundamentally dynamical in nature: their defining properties are causal, and their organisational and functional capacities are causally grounded. Among dynamical systems, an interesting and important sub-class are those that are autonomous, anticipative and adaptive (AAA). Living systems, intelligent systems, sophisticated robots and social systems belong to this class, and the use of these terms has recently spread rapidly through the scientific literature. Central to understanding these dynamical systems is their complicated organisation and their consequent capacities for re- and self-organisation. But there is at present no general analysis of these capacities or of the requisite organisation involved. We define what distinguishes AAA systems from other kinds of systems by characterising their central properties in a dynamically interpreted information theory.

1. Introduction

Both natural and engineered systems are fundamentally dynamical in nature: their defining properties are causal, and their organisational and functional capacities are causally grounded. Among dynamical systems, an interesting and important sub-class are those that are autonomous, anticipative and adaptive (AAA). Living systems, intelligent systems, sophisticated robots and social systems belong to this class, and the use of these terms has recently spread rapidly through the scientific literature. Central to understanding these dynamical systems is their complex organisation and their consequent capacities for re- and self-organisation. But there is at present no general analysis of these capacities or of the requisite complex organisation involved. We define what distinguishes AAA systems from other kinds of systems by characterising their central properties in a dynamically interpreted information theory.[1]

A satisfactory dynamical account of AAA capacity must bring together the resources of physical theory and process organisation theory into a unified theory. This presents a number of problems, not the least of which is formulating a dynamical account of process organisational notions, such as that of control.[2] Whereas

*This paper has reached its present formulation through joint discussion and it is impossible now to disentangle our individual contributions; however, one of us (JDC) has especially contributed the initial ideas for our treatment of information while the other (CAH) contributed initial ideas on systems (types, modularity). We also want to acknowledge valuable discussions with co-researcher Wayne Christensen (see references) and Bruce Penfold as well as Mark Bickhard, Jonathan D.H. Smith and Tim Smithers.

physical theory primarily involves the disposition and time evolution of energy and matter, control theory is more concerned with the organisation of interaction and information flow within a system (though its ultimate aim is controlling the disposition of matter and energy). The central problem is that a very small amount of information can alter the dynamical behaviour of a very large amount of energy and matter, as when a change of state of a bit of information in a computer can control a counterweighted dam gate to release megalitres of water. The controlling effort is minuscule compared to the effect, but it is not negligible: any control device must do some work to change its control state. It turns out that work is the common ground for informational and dynamical processes.

Most approaches to the problem since Descartes have been dualistic, treating the organisational design purely functionally (note 2) and then treating a functional design and its physical realisation separately, the only constraint being that the physical realisation can implement an appropriate functional design. Little attention is given to how functional requirements can be satisfied dynamically, in particular to the constraints that proper functioning places on the dynamics of the physical implementation.[3] These problems are seen as primarily technical, rather than central to the nature of functionality itself. This approach has the inevitable consequence of pushing back the origin of teleology into the mists of an originating mind. In robotics and AI, one is seduced into accepting purpose as an extension of the purposes of the human designers rather than as arising intrinsically in an appropriately organised system. In cognitive science, intentionality is presupposed by restricting psychology to the study of the "cognitively penetrable" [116] and focusing it on functional explanations [46]. In biology, teleology becomes an embarrassment. Nature, through its role in selection, not only tends to take over the traditional role of God (as was recognised, though controversially, soon after the publication of Darwin's *The Origin of Species*), it becomes difficult to even investigate the natural origins of purposive creatures.

Not all control theory ignores dynamical organisation. A number of authors[4] argue that consideration of dynamical processes can illuminate the minimal functional requirements for types of control problems. The outcome of this strategy is that many control problems turn out be much simpler than they would appear to be from a purely Cartesian computational perspective. At the same time, the strategy reveals the extent to which the Cartesian presupposition of the intentionality of the mental obscures a genuinely difficult problem: if relatively simple dynamical devices can solve apparently complex functional problems, what use are intentionality and symbolic computation? To answer this, we need to determine a class of interesting problems that either require intentionality or are advantageously solved by Cartesian reasoning.[5] We do not address this problem here but attempt to lay a principled foundation from which it can be addressed, along with other proper questions concerning the nature of living and intelligent organisational processes, by showing where, and on what basis, those capacities which ground them (namely AAA) fit into dynamical characterisations of systems.

Living systems are not passively independent, in the way a rock's crystalline structure is undisturbable by all but the most violent signals from its environment. Rather, they are vulnerable to disruption by impinging signals — storms, predation, cold ... — and constantly in need of replenishing their dissipating energy and order. The cohesive order of living systems must be actively regenerated by processes of various kinds (cellular reproduction, structural repair, energy supply, and so on). Their structural bonds have energies measured in electron volts, even fractions thereof, not the millions of electron volts that fix a rock into responseless stability. This explains why systems of this kind are adaptable, for unless they can constantly adapt to mitigate or compensate for disturbing signals they will be disrupted and, losing their cohesion, lose their identity as that sort of system.[6] This same vulnerability is the basis of their adaptability, since their internal delicacy makes them easily alterable, allowing them both to be sensitive to signals and to respond to these signals malleably and flexibly in order to regenerate themselves. Their responses to signals cannot be mostly passive, like those of a gas, nor largely uniform, like those of a crystal, but must so interrelate as to preserve the organised complexity that underwrites that very responsiveness and adaptability. This active independence, their characteristic organisational property, we will call their autonomy.[7]

Systems of this kind have many subsidiary organisational properties, i.e. system properties that are preserved by the open cycle of interaction with the environment. The Cheetah's capacity for the rapid chase is a complex neuro-optico-muscular property which is regenerated through the food successful chases provide, and through its very exercise (itself a neural entrenchment capacity arising from more basic properties of neurones ultimately regenerated through their support of entrenchable skills like chasing). Each such capacity adds to the adaptive fitness of its system. Indeed, autonomous systems are intrinsically organisationally global: because their capacities, i.e. their organised processes of interaction that ground their functional properties, must be so integrated that they are able to actively regenerate themselves, their overall functionality can not be grounded in a mere aggregate of independent processes but requires that distinctive global process integration that alone insures regeneration of the whole as a joint interactive consequence across all their interrelated process cycles. Process control in such integrated systems is typically complex, acting across many different dynamical timescales (cf. feeding cycles with moulting cycles) and requiring coherent activation and modulation (including nested sequencing) of subsidiary processes (cf. feeding following hunting, with optical and motor sequences nested inside hunting); this requires a system capacity for adaptability. The same requirement is reinforced by the need for effective response to changing environments. The selective advantage of being able to relevantly improve various functionalities completes the grounds for the importance of, not just adaptation, but adaptability.

Such autonomous, adaptive (adapted and adaptable) systems are intrinsically anticipative. Their capacities imply that their actions anticipate responses that will support autonomy, including those capacities. Behaviour, in particular, is fed forward in anticipation of receiving desired response signals from environment and self. Hunting behaviour, for example, is feed forward action in anticipation of receiving subsequent hunger satiation signals. Anticipative feedforward is fundamental to all self-controlling systems, to intelligent systems in particular; it combines with error-corrective feedback to deliver powerful learning and response capabilities. And thus we arrive at the AAA systems we wish ultimately to characterise.

We suggest that the roots of intelligence lie in the complex organisational requirements of AAA systems, that intelligence is essentially an articulated refinement of AAA capacities. As just noted, AAA systems already display subtly integrated stimulus-response capacities subject to considerable anticipative, modifiable endogenous control; this is already to show some of the central hallmarks of intelligence. Once the dynamical nature of that internal constitution is properly appreciated, we contend, it will be clear how intelligence, and our kind of conceptual intelligence in particular, is a natural extension, a natural refinement, of it. It is not the purpose of this paper to carry out that task[8], here our prior aim is to lay the foundation for that understanding by developing a general and principled account of complex dynamical systems which shows what systems have to be like to be AAA.

A key to approaching our task here is Schrödinger's [125] *negentropy principle of information*. Information is a very abstract and powerful concept that can in principle represent all computational and formal relations, conditions and constraints quantitatively.[9] These formal properties of information present us with some hope that system organisation, including cognition, can be quantitatively represented within the language of information theory. The negentropy principle, on the other hand, gives a quantitative connection between information and thermodynamic state variables of physical systems. The result is a dynamical information theory (in distinction to the abstract formal information theories of mathematicians and communications theorists.) Although this goes some way towards reducing the gulf between sophisticated functional (especially cognitive) and dynamical characterisations, it does not tell us how to do it.

Schrödinger [125] also suggested that the defining characteristic of life was the capacity to feed off exergy (available energy) to maintain and produce negentropy (system orderedness). While this is correct, the negentropy principle alone cannot distinguish intelligent or even living systems from hurricanes, stream eddies, sunspots, or the universe itself. The key to understanding Schrödinger's claim lies in unpacking the phrase "maintain and produce"; specifically we propose that living systems are distinguished from other non-equilibrium phenomena at least by their AAA capacities, as well as by their ability (R) to biologically reproduce themselves, including their own regenerative capacity. (A sine qua non of these systems is that AAAR is itself an invariant of both autonomous regeneration and biological reproduction.[10]) But it turns out that these properties are also intimately tied to concepts of complexity and self-organisation, which are in turn amenable to an information theoretic treatment. The rest of this article is an extended investigation into selected aspects of this treatment.

2. Organised Complexity Is Essential, Self-Organisation Is Characteristic

The AAAR systems with which we are most familiar — living systems, especially ourselves — are all very complicated. It would be nice if we could construct a model of a simple system that possessed these capacities and thereby obtain a better understanding of their essentials. Unfortunately, there are reasons (already suggested) to believe that the systems we wish to understand are inherently complicated, and that to understand such systems we need to understand the relevant kind of complicatedness. Our strategy will be to consider increasingly complicated kinds of systems, examining their limitations to see what additional features are required.

In classical physics, the tractability of three types of systems has meant that they have received most of the study. These systems are 1) single particle and conservative, decomposable (linearisable) multi-particle systems, 2) statistically specified systems at or near to equilibrium (e.g. gases, fluids), and 3) sufficiently well constrained but non-linearisable multi-particle systems, e.g. many machines and some electromagnetic systems. What each of these three kinds of systems share in common is the existence of analytic solutions for their dynamics, or convergent higher order additive approximations to these. In one sense that is why these systems are tractable, but this response does not illuminate the physical basis for their tractability. To do that we need to distinguish two different dimensions to tractability, often conflated or confused, complexity and organisation. Essentially, complexity refers to the number of independent pieces of information needed to specify a system (whether the specification is from an internal or external perspective), while organisation characterises the extent of the interrelations among the components of the system in terms of their number, scope and dynamics (degree of non-linearity). Tractable systems reduce the burden of accurately modelling their dynamics (and hence functioning) by exhibiting only low values in one or both of these two dimensions. Type 1 systems are uncomplex and unorganised (at most additive composition); type 2 systems are complex but unorganised; type 3 systems are uncomplex but organised. Complexity and organisation are relative notions in the sense that they are matters of degree, but they are not arbitrary. To a first crude approximation we can classify systems under the two dichotomies: complex versus uncomplex (simple), and organised versus unorganised (random). Our classification depends to some degree on physical scale: At a large scale organisational details or complexity at smaller scales may be irrelevant. Like the relativity of strength of complexity and organisation, their relativity of scale, although inconvenient, is not arbitrary.

The three types of systems tractable to classical physics fill up only three of the possibilities created by the complexity and organisation dichotomies. The fourth possibility comprises complex but internally organised systems. It is perhaps obvious that known living systems, especially cognitive and social systems, fit into this fourth category. Unlike systems of types 1 and 3, whose organisation, if any, is fully determined by their initial internal and boundary conditions, some type 4 systems can produce new organisation through time. And, unlike type 2 systems, whose organisation, if any, is entirely imposed by boundary conditions, type 4 systems contribute internally to regenerating their organisation and, where it increases, to increasing it. Some natural non-living type 4 systems are weather systems, stream eddies, and solar systems.

In properly understanding this classification it is important to note that the strength and scale relativities of complexity and organisation allow type 4 systems to mimic type 1, 2 or 3 systems under certain conditions. Specifically: A) over short time scales, except for type 4 systems near critical points in system phase space, behaviour is relatively linear. Only over longer time scales or near critical points do characteristic type 4 properties reveal themselves. For example, the revolution and rotation periods of Mercury are in a 2:3 ratio. On the small scale, this is explained by the analysis of planetary dynamics using gravitational theory with additive higher order perturbations. On the large scale, however, the question arises, why this harmony rather than none or the more expected 1:1 ratio? The answer requires understanding how order can arise through dissipative processes in systems with multiple attractor basins, a characteristic type 4 problem. B) On both small and large scales relative to system size, organisational properties are relatively unimportant. For example, the complex dynamics of stellar interactions can be treated largely statistically at the scale of galaxies, and as a type 1 system for binary stars; however the dynamics of globular clusters, especially their stability conditions, require type 4 analysis. C) The organisational properties of most systems are relatively insensitive both to interactions with much more strongly and much more weakly cohesive systems. For example, the human body has immense internal organisation crucial to its nature and yet, respectively, interactions between the human body and the air can be treated largely through particle mechanics and statistical fluid dynamics, and a human body falling onto rocks can be treated in terms of the theories of fluid and rigid bodies. Interactions between human bodies and food, however, cannot be completely understood without understanding the organisation and anticipatory capacities of human bodies, which we will characterise below as a type 4 problem.

Current computers are an interesting case because they are made of highly redundant components put together in a modular way. Their high redundancy and moderate physical organisation makes them type 3 systems, but they are capable of running highly complex and organised programs, and have been used to approximately model typical type 4 processes. Although their general principles of operation are very simple, they appear complex to many users because of the complexly organised initial and boundary conditions imposed by their programs. In contrast, the operation of living systems is of at least the same order (and usually higher) of organised complexity as their boundary and initial conditions. If computers (and computer guided robots) are to be capable of type 4 behaviour, their programming will need to be quite a bit different than it is now, allowing a high degree of self-modification not under the control of initial constraints.[11] Physically, then, computers are type 3 systems, but with the addition of suitably designed programs they can show some of the functional characteristics of type 4 systems; however they are still far from being paradigmatic type 4 systems even under these conditions, their sometimes unexpected and novel behaviour notwithstanding, because their dynamics is highly constrained (largely to on-off behaviour).

Because type 1, 2 and 3 systems are much more tractable than type 4 systems, and both scientists and engineers understandably tend to work with what is tractable, there is a natural tendency to subsume type 4 systems under the tractable classes using various scales and approximations, and ignore the conditions under which such systems are not tractable. This natural tendency should be resisted, because it leads to neglect of characteristic type 4 system properties which are among the most interesting and practically important, if most challenging, in nature.

Ingarden et al. [78] usefully distinguish between microscopic, macroscopic and mesoscopic systems. Most often these are understood relative to human interests (e.g. relative to human scale), but the ideas can be applied relative to any natural cohesive scale. For example, galaxies are macroscopically cohesive relative to stars, microscopically cohesive relative to the entire cosmos, and mesoscopically cohesive on the spatial and temporal scales of galaxies themselves. Ingarden et al. suggest that the mesoscopic domain introduces a new set of issues, those for which conservative classical mechanical and equilibrium methods (together, Hamiltonian methods) are not well suited. We suggest that it is definitive of type 4 systems that their mesoscopic domain properties can not be reduced to either their microscopic or their macroscopic properties without significant loss of explanatory value. Organisation requires a scale of the same order as the organised system, so the appropriate scale for understanding organised systems is mesoscopic. In the case of type 3 systems, which are organised but not complex, the organisation can be understood as the sum, or limit of, expansion of small scale dynamical properties, or else as the analytical product of, or limit on, reduction of the scale of large scale functional properties. Current computers, for example, can be understood in both these senses, in terms of the sum of the dynamical contributions of their physical components (their highly modular construction insures this), and in terms of their functional role in human scale applications (typical contemporary well-behaved program design reduces problems to highly modular and predictable components). Type 4 systems, by definition, require a more holistic, nonreductive approach.

While type 4 systems may contribute to regenerating their own organisation, this is compatible with their organisation actually decreasing, e.g. when dissipation overwhelms the regeneration. Hurricanes disperse and creatures die. Even systems that can increase their organisation must eventually cease doing so when they have exhausted their internal organisational space and/or the capacity of their environment to supply new order to them. These are processes of *senescence* [119, 120]. We shall label the first, senescent decline, and the second, senescent fixation. Non-senescent systems that are increasing their organisation we will call *immature*. Systems whose organisation is constant divide into two, non-dissipative systems will be held at their organisational ceilings, so in senescent fixation, while dissipative systems will be in a steady state (a dynamic equilibrium) and we will call them stably non-senescent, or *mature*. Natural systems typically pass from rapid growth in the immature phase, to stable non-senescence at maturity, and then to senescent decline and finally extinction (as that type 4 system). Of course this can occur at different rates in different parts or sub-systems of a system; e.g. in Alzheimer's disease a person's neurological functioning declines while their cardio-vascular functioning typically continues, and the reverse case is also common.

Consider now the processes which tend to promote organisation, the *organising* processes. These can be of two kinds, those where no new macroscopic constraints are formed and those where new macroscopic constraints emerge (expressed either as structures or structured processes). The former are essentially passive filtering, sorting and re-arranging processes, and we will call them *re-organisation* processes, in contrast to the latter *self-organisation* processes. To facilitate comparison we consider an example of each. Coins are often sorted by being placed in a sorting box, above a series of graduated meshes ordered by mesh size (biggest size on top) and chosen so that each mesh size is intermediate between the corresponding coin sizes. The sorter is randomly jiggled horizontally and the coins are automatically sorted by size as each eventually falls to its appropriate mesh, the meshes acting as passive filters. The coins have acquired a small increase in organisation (von Fürster's order-from-noise principle, with gravity the ordering principle) but the coin+sorter system has simply re-organised. On the other hand consider a collection of randomly positioned molecules at rest at equilibrium and then again subject to jiggling so that they acquire random increases in momentum; but this time suppose that they interact exothermically above some energy threshold, so that first one pair bonds as a result of a particular fluctuation, releasing energy, and in consequence an increasing cascade of bonds are formed, with the ultimate result that the collection condenses into a rigid, regular hexagonal crystalline object, releasing energy to form a surrounding radiation field. The molecules have acquired a small increase in organisation (through the ordering of bonding and radiative dissipation), this time through self-organisation.

First, consider what these cases share in common. Both systems have available an ordering principle and both are subject to uncontrolled perturbations which are required to initiate the organising process. (Here "uncontrolled" means a perturbation whose variable features are unconstrained by existing system constraints, i.e. by system structures and processes.) These features are essential for either process — otherwise order would be created ex nihilo. Without an ordering principle there would be no coherent motion in either system, their elements would disperse at random like an uncontained gas. Without an input of perturbations, each internally (if idiosyncratically) ordered, there would be no external source of new order from which to generate the new system order and organisation.[12] For the same reason, in both cases these perturbations must be uncontrolled.[13] The point of their being random is only to ensure that sufficient possibilities are explored to make it probable that at least one them will contain order that can initiate system ordering in response. This requires, again in both cases, that the system dynamics (in particular the ordering principle) must be such that the order in sufficiently many inputs is effectively retained (not "washed out" by subsequent fluctuations) and perhaps even amplified. In the coin case gravity transforms the order in a horizontal fluctuation that brings a coin into an appropriate correlation with a mesh (namely, over a larger hole) into downward motion and these motions are retained and concatenated into sorter-scale order. In the crystal case bonding force transforms the order in motion correlation into retained bonding which concatenates into a crystal and also amplifies it radiatively to initiate other bondings, ultimately retaining it in the final radiation field. In both cases the systems pass through a symmetry-breaking transition which corresponds to the creation of order or physical information within them [38]. The coins begin randomly positioned, the correlation function among like coins identical in all directions, but end asymmetrically organised vertically; the molecules begin likewise but end hexagonally organised. In both cases the resulting system state changes some of its dynamical behaviour. The coins will now comprise a vertically organised sequence of masses, each of a distinctive size, density and total weight, where initially they comprised a single mass of yet different distinctive size, density and total weight, and these initial and final arrangements will each be characterised by a distinctive gravitational field, rotational inertia and so on, and hence interact distinctively with other objects. And similarly for the other case. Finally, in both cases the process is irreversible because of dissipation; the coins dissipate their falling kinetic energy as heat as they bounce to a halt on their mesh and the molecules dissipate their energy as release of radiation as they bond.[14] All of these features are essential to all organising processes — and thus none of them distinguishes between re- and self-organisation.

Consider then what distinguishes between the processes in the two cases. First, there is new cohesion formed in the crystal case but not in the coin case. The new constraints, respectively on the molecules in the crystal matrix and on the quanta in the radiation field, are each markedly different from those applying to the unbound molecules; the crystal molecules are now bound together cohesively into a rigid object while the radiation quanta can freely superpose, neither of which applies to the movement of the individual free molecules. By contrast there are no new cohesive constraints formed in the coin sorting case, both initially and finally they simply act under gravity. More generally, a self-organising process is distinguished by the production of order at a higher constraint level (known as higher order redundancy in the communications theory literature) or the promotion of order from one constraint level to another. Second, and correlatively, neither the laws by which coins are governed nor those governing the sorter will have changed in the process, but in the other case the two emergent objects — crystal and radiation field — obey correspondingly different laws from those for free molecules. In the coin case the coin+sorter system uses its existing dynamical laws, in interaction with its environment, to reorganise itself without adding any new kinds of internal constraints and hence without changing its dynamical form. By contrast, during its organising process the change in laws across the crystal-forming process is a change in dynamical form. Moreover, third, the change in form is quite specific: whereas in its initial state all the system components (the molecules) could in principle interact with one another so as to form correlated movements, typically time-varying, in the final state the movements of the molecules are invariantly, repetitively correlated and movements of individual molecules are not correlated or correlatable with the movements of quanta in the radiation field, even when the two interact. From a single interactive dynamics the dynamics has bifurcated into a fixed cohesive dynamics plus a collection of fluctuations uncorrelated with it. By contrast, there is no such bifurcation in the coin case. Thus, fourth, in the re-organisation case there is no phase transition and in the self-organisation case there is. Fifth, the result of the new cohesion formation is the construction of a new system level which filters out fluctuations on smaller scales. In the crystal case movement fluctuations in individual molecules resulting from interaction with field quanta are not preserved but dissipated as re-radiated heat from the crystal, the crystal literally averaging across such individual fluctuations because of its cohesion. There is no corresponding formation of relatively macro scale filtering in the coin case.[15]

Finally, sixth, it follows that in the case of self-organisation new kinds of possibilities will emerge because of the new cohesive level created. Thus the system will be characterised by increasing degrees of freedom in the relevant respects. These may be wholly new; the collection of separated molecules has no degrees of freedom as a rigid body (thus for rigid rotation, sliding, strain and stress — e.g. when supporting weights, and so on) or as an electromagnetic wave conductor (e.g. as a polariser) while the crystal does. (The constraints are enabling, see [107, 65].) Of course other possibilities will also have been surrendered. The collection of molecules can disperse, stream and eddy, collect in odd shapes and so on. (Constraints are also always disabling.) Thus there will be a sharp shift in degrees of freedom, a characteristic of phase transitions. In the re-organisation case there will be no such sharp shift since the dynamical form remains invariant, there can only be gradual quantitative alteration of existing degrees of freedom.[16] These contrasts generalise to all cases. All the features of self-organisation are present, e.g., in the paradigm cases of Bénard cell formation, slime mould aggregation and others and are clearly absent from all the cases of re-arranging, sorting and filtering that leave the original elements and their interrelations unchanged. To sheet home this often mis-understood distinction, consider the following subtlely different cases. First, a standard learning system, e.g. a neural net equipped with some error correction process, learns to discriminate environmental stimuli and provide a discriminated response. Second, consider the same example, but now with the net weights so arranged that if they fall in value below some threshold they pass irreversibly to zero, i.e. to effective disconnection.[17] In both cases the stimuli are uncontrolled perturbations for the net, environmental order in the stimuli is incorporated into system order and learning will correspond to increased system organisation. However, in the first case the laws by which the learning system works, e.g. neural net node and error correction rules, will not have altered and all net elements remain connected. The interactive system+environment super-system evolves from its initial condition under its unchanging dynamical laws and environmental information is simply transferred or copied into the system. In the second case, by contrast, the collapse of connections establishes new net-scale constraints which future fluctuations in net values cannot disturb, so that this new constraint formation acts as a macro level filter; the net changes its dynamical form and so passes through a phase transition. In short, in the first case we have learning through re-organisation and in the second case we have learning through self-organisation.[18]

Since in all realistic type 4 systems internal interactions not only support creation and maintenance of organisation but also degrade energy and order (second law of thermodynamics), these systems need to be constantly replenished by an input of ordered energy. Their structures and processes are sustained far from thermodynamic equilibrium by this flow. The river's standing waves, for example, are sustained by the flow of water downhill. Where fluctuations are fixated as new structure it will be this flow which sustains the new dynamic equilibrium. And where micro fluctuations are amplified to macro scale, it will be this flow which ultimately supplies the energy and order to do so. Thus type 4 systems are of nomic necessity far from equilibrium dissipative systems driven and stabilised by external ordered energy flows [114, 115, 105] and exhibit spatio-temporally distributed (i.e. "global", or "holistic") constraints on their behaviour.[19]

We know that these conditions govern the physiological processes that underlie typical living systems, all of which are AAA. Conversely, each of the AAA capacities requires type 4 endogenous organised complexity. Process control in autonomous systems, we noted, is typically complex, acting across many different dynamical timescales and requiring coherent activation and modulation (including nested sequencing) of subsidiary processes, which in turn requires complex endogenous organisation. Anticipatory systems must be complex and organised to be able to respond selectively and differentially to stimuli so as to likely produce a state of their environment that would otherwise be unlikely (roughly, that state most advantageous to the support of their autonomy). And adaptiveness, especially adaptability, also requires anticipative control over internal processes and behaviour, again requiring organised complexity. Thus the confirming evidence for the idea that organised complexity is required by AAA systems is strong.[20]

Type 4 systems are not maximally complex, nor are they maximally ordered. All other things being equal, ideal gases and perfect crystals respectively take the complexity and order kudos. Complexity and order are necessary to type 4 systems but not in themselves desirable. Complexity per se is a liability for highly organised, integrated systems because it adds to their endogenous co-ordination costs. Order per se is not desirable because the strongest forms of order are too constraining to be compatible with the kind of organised flexibility a type 4 system, especially a AAA system, must show to remain viable. To understand type 4 systems, we need instead to understand the significance of their kind of organisation and organising processes, especially self-organisation.

Another significant characteristic of type 4 systems has both a diachronic, or generative aspect, and a synchronic, or structural aspect. The AAA systems with which we are familiar have all evolved from simpler forms, and regenerate themselves from simpler, inherited material. Various evolutionary theorists have suggested mechanisms by which historical developments constrain further evolution and development. The process has been called *coevolutionary adaptation* [47], *canalization* [135], *generative entrenchment* [139, 13], and the self production of *historical constraints* [13], depending on theoretical framework. Similarly, physiological development involves a progressive constraining of later development, especially in maturity and senescence [119, 42]. And the same idea is common in theories of cognitive development.[21]

Much of the language used to describe historical constraints uses depth metaphors: levels, entrenchment, canalization, stages. Earlier historical constraints are buried deeper, and are harder to change without being lethal. This is reflected structurally in the way that *some* earlier constraints create complex interdependencies in later developments. For example, the surface structure of an organism, its *phenotype*, is constrained not only by the need to make the various parts fit together, but also by the need to generate the phenotype from superficially simpler forms (its *genotype* and basic metabolism). The same chemicals, for example hormones and neurotransmitters, can play vastly different roles in different cells and organs, but the overall regulation is inevitably interconnected. This sort of deeply co-ordinated structure is quickly recognised as "organic".

We emphasise that in all these cases the cohesive constraints are dynamically real, they arise from the actual dynamical interactions of the system components, and once formed they genuinely constrain the behaviours of those components. The molecular interactions that make for a rigid, non-conducting wooden table are dynamically different from those that make for a malleable conducting copper wire and so place different constraints on their component nuclei and electrons. These constraints are neither mere patterns, epiphenomenal, or observer relative.[22] Moreover, they deliver the kind of organised complexity that prevents our ever obtaining simple models of these systems. The kinds of changes characterising self-organisation, e.g., where there is an inter-play between relatively micro level dynamics and relatively macro filtering formation, presents cases of simultaneous "bottom up" and "top down" causal influences (better: interactive constraint formation) and we as yet do not have, and may never have, satisfying mathematical models of such processes.[23] We turn instead to the development of analyses that throw useful light on these inherently complicated, but ubiquitous and important, systems.

3. Informational Complexity

To understand the kind of organised complexity required for AAA systems requires a common formal language that can make rigorous sense of the notions of complexity, organisation, depth, entrenchment, canalization and constraint. None of these notions is in itself dynamical; they are at best abstractions from their underlying dynamics. The language must have a dynamical interpretation, i.e. a clear definition of non-dynamical notions in terms of dynamical concepts. Ideally, the definitions should allow us to measure, or determine from measurable quantities, the amount of complexity, organisation, depth, and so on. The only language that is both broad enough and rigorous enough to do the job is information theory.

Complexity has proven difficult to define. Different investigators, even in the same fields, use different notions.[24] The Latin word means "to mutually entwine or pleat or weave together". In the clothing industry one fold (e.g. in a pleat) is a *simplex*, while multiple folds are a *complex*. The most fundamental type of complexity is *informational complexity*. It is fundamental in the sense that anything that is complex in any other way must also be informationally complex. A complex object requires more information to specify than a simple one. Even the sartorial origins of the word illustrate this relation: a complex pleat requires more information to specify than a simplex, one must specify at least that the folds are in a certain multiple, so a repeat specification is required in addition to the "produce fold" specifications. Further information might be required to specify any differences among the folds, and their relations to each other.

The size of a structure (defined, for example, in terms of the sum of the number of its nodes or elements that have values and the number of the particular value relations among nodes) is not a good guide to its informational complexity. Two structures of the same size or made from the same components might have very different informational complexities if one of the structures is much more regular than the other. For example, a frame cube and a spatial structure composed of eight irregularly placed nodes with straight line connections between each node may encompass the same volume with the same number of components, but the regularity of the cube reduces the amount of information required to specify it. This information reduction results from the mutual constraints on values in the system implied by the regularities in the cube — all the sides, angles and nodes must be the same. This redundancy reduces the amount of information required in a program that draws the cube over that required by a program that draws the arbitrary eight node shape. Similarly, a sequence of 32 '7's requires a shorter program to produce (namely one specifying 5 doublings of an initial output of '7') than does an arbitrary sequence of decimal digits. To take a less obvious case, any specific sequence of digits in the expansion of the transcendental number $\pi = 3.14159\ldots$ can be produced with a relatively short program, despite the apparent randomness of expansions of π. The information required to unambiguously describe certain types of structures can be compressed due to the redundant information they contain; other structures can not be so compressed. This is a property of the constraints contained in the structures, not directly of any particular description of the structures, or language used for description.

The informational complexity of a structure s can be given a precise, mathematical definition: Let s be mapped isomorphically onto some binary string σ_s (so that s and only s can be recovered from the inverse mapping), then the informational complexity of s is the length in bits of the shortest self-delimiting computer program that produces σ_s, minus any computational overhead required to run the program, i.e. $C_I = \text{length}\,(\sigma_s) - O(1)$. The first (positive) part of this measure is often called *algorithmic complexity*, though it is also called computational complexity, or Kolmogorov/Chaitin complexity. The second part of the measure, $O(1)$, is a constant (order of magnitude 1) representing the computational overhead required to produce the string σ_s.[25] This is the complexity of the program that computes σ_s. It is machine dependent, but can be reduced to an arbitrarily small value, mitigating the machine dependence.[26] We deduct it to define the informational complexity because we want a machine independent measure that is directly numerically comparable to Shannon information[27], permitting a rigorous formal identification of algorithmic complexity and combinatorial and probabilistic measures of information. A nearly complete proof of the interchangeability of the different concepts of information is given in [78] (pp. 25ff — the proof lacks an explicit derivation of the algorithmic formulation of information). They argue, and we agree, that there is only one underlying notion of syntactic or formal information. We hold that this notion is ultimately based in the capacity to make distinctions (see next section on work capacity, especially the sorting example).

All non-computable strings are algorithmically random [93]. They cannot be compressed, by definition; so they contain no detectable overall order, and cannot be distinguished from random strings by any effective statistical test. This notion of randomness can be generalised to finite strings with the notion of *effective randomness*: a string is effectively random if it cannot be compressed. Random strings do not contain information in earlier parts of the sequence that determines in any way later members of the sequence (or else they could be compressed).[28] Thus any system whose trajectory cannot be specified in a way that can be compressed is dynamically disorganised. It cannot anticipate, control, or show any other like features, except coincidentally.

4. The Negentropy Principle of Information

To connect information theory to dynamics, it is useful to define the notions of order and disorder in a physical system in terms of informational complexity. The concept of disorder is connected to the concept of entropy, which has its origins in thermodynamics, but is now largely explained via statistical mechanics. The statistical notion of entropy has allowed the extension of the concept in a number of directions, directions that do not always sit happily with each other. In particular, the entropy in mathematical communications theory [126], identified with information, should not be confused with physical entropy.[29] Incompatibilities between formal mathematical conceptions of entropy and the thermodynamic entropy of physics have the potential to cause much confusion over what applications of the concepts of entropy and information are proper (e.g. [138, 12]).[30] One way to control such problems is to restrict the use of the concepts of entropy and information, respectively, to the thermodynamic and communications realm [48], keeping their quantitative budgets separate, so that no conflict can arise. This move, though, can obscure interesting relations between the entropy and information budgets, and also prevents an illuminating unification of their respective theories, which might permit extension of the joint concepts to new areas.[31]

We prefer to adopt the interpretive heuristic known as the Negentropy Principle of Information (NPI), according to which the information in a specific state of a physical system is a measure of the capacity of the system in that state to do work [125, 11], where work is defined as the application of a non-inertial force in a specific direction, through a specific distance.[32] Through this connection with work, NPI ties information, and hence complexity and order as well as the concepts of constraint, sorting and selection required to understand regulation and control, to dynamical concepts (e.g., see the sorting and steam engine examples below, and the discussion of constraints, including regulation and control, in Section 8). NPI implies that *physical information* [11] has the opposite sign to physical entropy[33], and represents the difference between the maximal possible entropy of the system (its entropy at equilibrium with its environment, assumed otherwise unchanged, after all cohesive constraints internal to the system have been removed) and the actual entropy, i.e.

$$I_P = H_{\mathrm{MAX}}\Big|_{E=\mathrm{const}} - H_{\mathrm{ACT}}\Big|_{\substack{E=\mathrm{const}\\ C=\mathrm{const}}},$$

E the environment, C the constraints. The actual entropy, H_{ACT}, is a specific physical value that can in principle be measured directly [1], while the maximal entropy, H_{MAX}, of the system is also unique, since it is a fundamental principle of equilibrium thermodynamics that the order of removal of constraints does not affect the value of the state variables at equilibrium [82]. This implies that the equilibrium state contains no trace of the history of the system, but is determined entirely by synchronic boundary conditions. Physical information, then, is a unique measure of the amount of form, order or regularity in the physical system. Its value is non-zero only if the system is not at equilibrium with its environment. It is fundamentally a measure of the deviation of the system from that equilibrium. It is important to remember that the Negentropy Principle of Information (NPI) is a heuristic physical principle, not a formal or operational definition and, given the current proliferation formalisms for entropy and information, it needs to be interpreted as appropriate for a given formalism and for a given physical system and its environment.[34]

On the other hand, NPI has some of the formal properties of a definition, since it determines how terms like entropy and information are to be used in a physical context. As in mathematics, central definitions in empirical theory should be supported with an existence proof. This is done by showing that violating the definition would violate known observations [94, pp.264ff]. If we assume NPI, then reliable production or reproduction of one bit of information requires a degradation of at least $kT \ln 2$ exergy (available energy), where k is Boltzmann's constant in a purely numerical form [11], and T is temperature measured in energy units. This relation must hold, or Maxwell's demon will come to haunt us, and the Second Law of Thermodynamics will come tumbling down.[35] NPI is empirically justified; we know, for example, that violation of NPI, which would amount to using information to reduce the entropy of an isolated system, violates our most common experiences. NPI implies that a bit of information can be identified with the minute but non-negligible physical value $k \ln 2$ and that its transfer from one system or part of a system to another will require the *transfer* of at least $kT \ln 2$ exergy (i.e. entropy change $= -$(information change) $=$ heat change$/T$, $kT \ln 2/T$ in this case, whence heat change $=$ maximum work done $= -T$(information change); see [11]). This gives us a quantitative physical measure of form that is directly related to exergy and entropy, the central concepts in non-equilibrium processes. These relations allow us to study complexity changes in physical processes, and permit principled extensions of the concepts of entropy and information.

To take a simple example, imagine that we start with a container of m "red" and n "white" molecules in an ideal gas at equilibrium, S_0, and it ends up in a state, S_1, in which all the red molecules are on the right side of the container, and the white molecules are on the left side, so that we could move a frictionless screen into the container to completely separate the red and white molecules without doing any additional work. The entropy of S_0 is $-\sum P_0 k \ln P_0$, and the entropy of S_1 is $-\sum P_1 k \ln P_1$, where P_0 is the inverse of the number of complexions in the initial state, and P_1 is the inverse of the number of complexions in the final state. Simplifying again, assume the $m = n = 1$.[36] Then the entropy of the final state is obviously 0, since there is only one possibility, in which the red molecule is on the right, and the white molecule is on the left, so $P_1 = 1$. The entropy of the initial state is higher: both molecules can be either on the right or the left, or there can be a red on the left or a red on the right, giving four distinct possibilities, and $P_0 = 1/4$. If we know that the system is in S_1, we have 2 bits more information than if we knew merely that it was in S_0. For example, we might have the information that no two molecules are on the same side, and that a red molecule is on the right, requiring two binary discriminations. To slide the screen in at an appropriate time, we need the information that the system is in S_1, i.e. we need the information difference between S_0 and S_1. This is exactly equivalent to the minimum entropy produced in a physical process that moves the system from S_0 to S_1, as can be seen by setting k to 1, and using base 2 logarithms to get the entropy in bits. To move the system from S_0 to S_1, then, requires at least $2T$ work. This is a very small amount; the actual work input would be larger to cover any energy stored and/or dissipated. Alternatively, a system in S_1 can do at most $2T$ work before it has dissipated all its available energy from this source. Putting this in other words, the system can make at most two binary distinctions, as can be seen by reversing the process.[37] These two bits measure the maximal controlling potential of the system: implemented as a controller, controlling either itself or another system, the system could function as at most two binary switches. Calculating the physical information for each case from the definition above, $I_P(S_0) = 0$, while $I_P(S_1) = 2$. As it should, the difference gives us the amount of information lost or gained in going from one state to the other. A number of years ago it was confirmed that the entropy production of the kidneys above what could be attributed to the basal metabolism of its cells, could be attributed to the entropy produced in sorting molecules for elimination. Presumably, more subtle measurements would confirm a physical realisation of our example.

The relations between information and work capacity are somewhat subtle, since they involve the correct application of NPI, which is not yet a canonical part of physics. The physical information in a given system state, its capacity to do work, breaks into two components, one which is not constrained by the cohesion in the system, and one which is. The former, called *intropy*, $\imath$, is defined by $\Delta\imath = \Delta(\text{exergy})/T$, so that $\int T\Delta\imath$ measures the available energy to do work, while the latter, called enformation, ε, measures the structural constraints internal to the system that can guide energy to do work.[38] Intropy measures the ordered energy which is not controlled by cohesive system processes, i.e. by system laws, it is unconstrained and hence free to do work. For this reason, though ordered, both intropy and exergy are system statistical properties in this sense: their condition cannot be computed from the cohesive or constrained system state, the cohesive state information determines the micro state underlying the intropy only up to an ensemble of intropy-equivalent micro-states. There is another system statistical quantity, entropy, e, but it is completely disordered or random, it cannot be finitely computed from any finite system information.[39] Entropy is expressed by equiprobable states, hence appears as heat and has no capacity to do work, relative to the system constraints for which it is defined; $\Delta e = \Delta Q/T$, where Q is heat, and $\int T\Delta e$ measures heat. Enformation is required for work to be accomplished, since unguided energy cannot do work.[40] Intropy is required for work in dissipative systems, to balance dissipation (e production).

Consider, for example, a system S with heat Q as its only unconstrained energy. If S is at equilibrium then the only enformation is the existence of a system temperature (not that it is of some specific value T), for only that follows from the system constraints, and $\imath = 0$ and Q is entropic since Q cannot do work on S. If S nomicly maintains an internal temperature gradient G then G is enformation for S since it cannot be released to do work without first altering the cohesive structures of S. If G is unconstrained by S then G expresses intropy in S since G is an ordering of the heat energy and work can be done in S because of G. (In fact S will dissipate G, creating entropy, until equilibrium is reached.) Further, note that if S, even if at internal equilibrium with $G = 0$, is made part of a larger system S' where it is in contact with another sub-system P of S' at a lower temperature, then there is now a new temperature gradient G' unconstrained by S' so S will do work on P with heat flowing between them until equilibrium is reached ($G' = 0$) at some intermediate temperature; hence G is intropic in S' even though S has no intropy and S's temperature, which serves in part to determine G', is enformation in S.[41] These analyses carry over to all other physical forms of energy.

The main difference between intropy and enformation is the spatial and temporal scale of the dynamical processes that underlie them.[42] The dynamics underlying intropy have a scale smaller than that of the whole system, and involve no long term or spatially extended constraints except those that govern the system as a whole, which in turn constitute the system enformation. The intropy of a system S is by definition equal to the difference between S's actual entropy and its maximal entropy when exergy has been fully dissipated (given enformation invariant, i.e. S's constraints remaining unchanged, and environment invariant); so,

$$\imath = I_\imath = H_{\mathrm{MAX}}(S)\Big|_{\substack{E=\mathrm{const}\\ C=\mathrm{const}}} - H_{\mathrm{ACT}}(S)\Big|_{\substack{E=\mathrm{const}\\ C=\mathrm{const}}} .$$

The enformation is just the additional information equal to the difference between $H_{\mathrm{ACT}}(S)$ and the entropy of the set of system components that result when the constraints on S are fully dissipated and S comes to equilibrium with its environment (assumed to remain otherwise invariant);

$$\varepsilon = I_\varepsilon = H_{\mathrm{MAX}}(S)\Big|_{E=\mathrm{const}} - H_{\mathrm{MAX}}(S)\Big|_{\substack{E=\mathrm{const}\\ C=\mathrm{const}}} .$$

Note that

$$I_P(S) = \imath + \varepsilon = H_{\mathrm{MAX}}(S)\Big|_{E=\mathrm{const}} - H_{\mathrm{ACT}}(S)\Big|_{\substack{E=\mathrm{const}\\ C=\mathrm{const}}} .$$

as required by NPI. This is perhaps more clear with an example. A steam engine has an intropy determined by the thermodynamic potential generated in its steam generator, due to the temperature and pressure differences between the generator and the condenser. Unless exergy is applied to the generator, the intropy drops as the engine does work, and the generator and condenser temperatures and pressures gradually equilibrate with each other. The enformation of the engine is its structural design, which guides the steam and the piston the steam pushes to do work. The design confines the steam and the piston in a regular way over time and place. If the engine rusts into unrecoverable waste, its enformation is completely gone (as is its intropy, which can no longer be contained), and it has become one with its supersystem, i.e. its surroundings. Such is life.

As noted, NPI allows us to divide a physical system into a regular, ordered part, represented by the physical information of the system, i.e. $I_P = \imath + \varepsilon = I_\imath + I_\varepsilon$, and a random, disordered part, represented by the system entropy e. The *orderedness* of the system is its information content divided by the equilibrium (i.e. maximal) entropy, i.e.; $O = I_P/H_{\text{MAX}}$, while the disorderedness is the actual entropy divided by the equilibrium entropy, i.e. $D = H_{\text{ACT}}/H_{\text{MAX}}$ [90, 89]; it follows from NPI that $O + D = 1$. The informational complexity of the information in the system, $C_I(I_P)$, is equal to the information required to distinguish the macrostate of the system from other macrostates of the system, and from those of all other systems made from the same components.[43] The mathematical relations between statistical entropy and algorithmic information [83, 84] ensure that $C_I(I_P) = H_{\text{MAX}} - H_{\text{ACT}}$, so $C_I(I_P) = I_P$. This is so since the physical information of a system determines its regularity and this regularity can be neither more nor less informationally complex than is required to specify the regularity. (The informational complexity of the disordered part is equal to the entropy of the system, i.e. $C_I(H_{\text{MAX}} - I_P) = C_I(H_{\text{ACT}}) = H_{\text{ACT}}$, $C_I(S) = C_I(I_P) + C_I(H_{\text{MAX}} - I_P) = C_I(H_{\text{MAX}}) = H_{\text{MAX}}$ and since $O = I_P/H_{\text{MAX}}$, the ordered content of $S = H_{\text{MAX}}O = I_P$ as required.) These identities allow us to use the resources of algorithmic complexity theory to discuss physical information, in particular to apply computation theory to the regularities of physical systems. This move has always been implicit in the use of deductive reasoning to make physical predictions, and should be non-controversial. The main achievement here is to explicitly tie together computational and dynamical reasoning within a common mathematical language (see also [6]).[44] As noted, the system intropy is an important resource, measuring system self-organisation potential.[45] The self-organisation potential is a measure of the room a system has to spontaneously generate new cohesive organisation or structure.[46]

5. Conservative Systems

We can now begin to discuss the information budget of dynamical interactions through the connection between information and physical action sanctioned by NPI. To best illuminate the inherently complicated AAA systems, we start with simple predictable systems and move to more complex ones as required.

It is convenient to classify all dynamical systems as conservative or dissipative. Conservative systems have H_{ACT} = constant. They are therefore closed systems, so that the sum of the potential and kinetic energy is a constant, they have a time-invariant Hamiltonian, and their dynamics are reversible [60]. In realistic systems, in which exergy eventually dissipates, this implies that I_P = constant, $\imath = 0$ (since any internal order that is not constrained has been dissipated already, or the system would not be conservative), and $I_P = \varepsilon$. Examples are equilibrium thermodynamic systems and classical particle systems. In the former case the enformation is determined by the properties of the external constraints on the system (e.g. the container of an ideal gas), while in the latter it is determined by the location of the system in phase space at any given time. In a conservative system, then, dynamical changes must involve shifts of enformation within the system that are permitted by the dynamical symmetry laws governing conservation of energy and momentum, i.e. system re- organisation only.

These constraints are severely restrictive. Conservative systems cannot support the sort of organised complexity we require. A complete representation of the microstate of an ideal gas at equilibrium, for example, is very complex algorithmically, but its disorder makes it an unsuitable candidate AAA system. Its regular properties are few (low I_P), and are entirely determined by the boundary conditions of the gas. Internally, an ideal gas is almost completely disordered (fluctuations can create but very temporary regions of order), and contains no useable information. Adaptation and anticipation (forms of work) require rule-conforming (regular) complex responses to diverse stimuli that a system with few regular properties is informationally incapable of martialling. Autonomy at the very least requires an internal control of regular activity, not merely boundary condition determination. So ideal gases, though their complete representation has a large C_I, fail on both counts to be autonomous and anticipatory.[47] All other type 2 or approximately type 2 systems would fail on the same grounds.

Perfect crystals are at the other, simple extreme. They are ordered, since their complete specification needs only the structure of the components together with a local concatenation routine. This also implies a low overall C_I, and a corresponding inability to show any of the AAA capacities: Perfect crystals are too regular, too strongly constrained. The entropy of a perfect crystal is 0, so $I_P = C_I$, and its order is complete, i.e. $O = 1$. A perfect crystal should be immune to deformation that would increase its C_I and lower its orderedness. This seems to ensure the independence of the crystal from boundary conditions, but this is not the kind of autonomy relevant to AAA systems, since it is precisely their organised interactions with their environment and their organised internal modifications to support those processes that expresses their AAA capacities. (The crystal is not internally self-regenerating and solves no problems at all, because it has none of the relevant kinds of capacities.) Thus perfect crystals, and like simple systems of classes 1 and 3, also fail as candidate AAA systems.

The enformation can be divided up into *conformational information*,[48] I_C, and *dynamical information*, I_D, where $I_C + I_D = I_P$ in the conservative case. Conformational information is contained in the position parameters of the Hamiltonian, while dynamical information is contained in the momenta and the relations between positions and momenta which represent the laws governing the system. Thus, the dynamical information includes the information contained in the laws governing the dynamics of the system.[49] It would be convenient to divide the enformation into separate parts representing the potential energy and the kinetic energy, but this is possible only if the potential energy is a function of positions alone (e.g. a system of two non- dissipative bodies in an inverse square force field). In that case, the information of the potential energy is the conformational information plus the information in the relevant force law(s). This division will always be possible in equilibrium and other linear systems, but general conservative systems are often nonlinear.[50]

Linear conservative systems are extremely simple in one sense: they are dynamically decomposable, allowing independent testing, analysis, engineering and control of subsystems. Near-to-linear systems can be approximated arbitrarily accurately with linear approximations and a series of linear corrections of higher order, as needed. This method is used extensively in physics and other areas of science. The main root exemplar is the analysis and prediction of the orbits of the planets where gravitational interaction is weak.[51] Despite the simplicity implied by decomposability, linear conservative systems can be informationally complex if ε is high ($\imath$, of course, is negligible), but this can only occur in a highly constrained structure which is nonetheless random (an amorphous solid, like glass, but more so). Such a structure would have high I_P due to its high ε, but, because of its randomness it would not be capable of organised interactions with its environment. These structures are extreme versions of aperiodic crystals, a class to which Schrödinger once suggested that DNA belongs.[52] Again, we see that complexity alone is not enough for organised behaviour.

Other forms of complex conservative systems are possible, if we allow nonlinear systems (at the expense of decomposability). These permit, indeed typically construct, rich organisational structures, i.e. their non-linearities impose constraints which produce correlational redundancies of many kinds, e.g. those of entrained oscillators and n-body planetary systems, both of which may show chaotic behaviour [140]. But much of the richness produced by non-linearity is most clearly and relevantly seen in non-conservative systems and we postpone its exploration until then, pausing here only to briefly place deterministic chaos within our framework.[53]

While perhaps currently the most notable non-linear systems are those that exhibit deterministic chaos, we have just noted that complex organisation does not require this. And we need to recall that even chaotic systems are deterministic and their specifications are typically highly compressible (e.g. to a few equations), so the appearance of complexity in the extensive manifestation of an algorithm may be misleading. Consider a Penrose tiling [108], which can be generated by a relatively simple recursive function, but which shows no obvious redundancy within any given region. It would be extremely difficult (if not impossible) computationally to recover with demonstrable certainty the generating function by sampling the conformation (pattern) of a Penrose tiling, so the pattern complexity or conformational information appears to be very high. This is an illusion, however, since all this information can be compressed to the generating function. Despite appearances, Penrose tilings are not very complex (i.e. they have low ε) though they are highly organised.[54] We now have a large catalogue of systems with these characteristics, which can be implemented on computers. It can be disarming that such apparently complex and even chaotic patterns can be generated by simple programs, since the simplicity of the programs proves the low ε of the patterns. We can imagine, however, chaotic systems that have a relatively high ε, are not locally redundant, but are still highly organised. Weather systems and living nervous systems quite possibly manifest this type of organisation — although these systems are not conservative, only partially chaotic, and perhaps not fully deterministic. So let us turn to considering organisation in general.

It might seem that conservative systems must have constant order, since their disorder is constant. Although this is true for equilibrium systems, for which the internal order is zero, structured systems can maintain constant disorder while order increases. This is possible (however physically unlikely), if the phase space of the system is expanding. H_{MAX} increases, while H_{ACT} remains constant, as required by conservation, thereby increasing I_P through an increase in intropy. Recall that availability and entropy, unlike energy, are not conserved; available work has increased because existing energy is not redistributed across new modes as fast as these modes are being created (e.g. between matter and radiation) and this ordered energy distribution is not constrained by the system, so some existing modes can now do work on others.[55] If the phase space expansion is a consequence of system dynamics, as appears to happen as a result of the spatial expansion of the universe, order can increase spontaneously as well. The surprise of the result that conservative systems can thus increase their order is mitigated by the fact that the intropy produced by expansion is extremely subject to dissipation, so we are unlikely to observe such an event. Nonetheless, if the expansion occurs faster than the system relaxes, some residual intropy will remain, and order will still increase. It turns out that this allows order and disorder to increase together, and permits an explanation of the origins of organisation. To be physically realistic, however, this takes us into the realm of non-conservative or dissipative systems.[56]

6. Organisation and Logical Depth

Organisation is the co-ordination or interdependence of parts or components, especially in support of vital functioning (OED). A living body is well organised when its organs so interrelate that the body as a whole can regeneratively maintain all its vital functions, i.e. is autonomous. A sporting team (football, relay, etc.) is properly organised when its members so interrelate that the team can perform its sporting function, and well organised when it performs well. Similarly for the parts of a machine. To be co-ordinated or interdependent is at least to be correlated in states and/or behaviour. But it is typically not to be identical in state or behaviour, for that produces only simple order (like the crystal, all of whose atoms stand in identical local relations to their neighbours); it is the different contributions of each component that gives the richness to organisation. Correlations entail descriptive redundancy; if A, B and C are correlated in respect X we may replace their independent description $\{A(X), B(X), C(X)\}$ with $\{A(X), R(A, B, C)\}$ and so on. So the search for a formal characterisation of organisation might profitably focus on understanding its specification in terms of redundancy.

Following [126] redundancy orders are determined by the minimal number of elements in which a redundancy can be detected, so the redundancy in a system (physically, its I_P) can then be decomposed into orders n based on the number of components, k_n required to detect the redundancy of order n.[57] Order 1 redundancy can be detected by examining elements of a system pairwise, whereas order n redundancy is dectectable over a minimum of $2n$ elements. Examples of low order redundancies are the simple repetitions of molecular arrangement characterising a perfect crystal and imposed by its molecular bonding and the requirement that being a word of English places on sequences of letters. Examples of related higher order redundancies are the long-range correlations in some crystal pattern fracturing produced by strain or the introduction of impurities and those imposed by being a sequence from a possible lost Shakespearian play or being a sequence of letters from a PhD thesis. In the higher order cases a much larger sample is required to recognise the appropriate property than is required to recognise the properties in the low redundancy cases. We can recognise a play as a sample of English with a much smaller sample than we can recognise it as a genuine work of Shakespeare. As the examples suggest, high order redundancy is also much harder to produce than low order redundancy, independently of the complexity of the sequences that contain them.

There is only a very loose constraint connecting high and low order redundancy: For fixed $I(p)$, high order redundancy implies less than maximal low order redundancy (a completely redundant system, e.g. a row of 1's or a perfect crystal, could show no high order redundancy). Otherwise high and low order redundancy are found to vary widely, effectively mutually independently, across real systems (which also show wide variations in $I(p)$). First, while large low order redundancy may constrain high order redundancy, there are systems with substantial measures of both. Consider, for example, the dynamics of a typical digital computer running a coherent program, e.g. drawing a complicated picture; its dynamics are locally redundant, since the local transitions are highly constrained, but the overall program coherency provides considerable higher order redundancy. Similarly, there is sufficient local low- order redundancy at cellular level in living creatures to maintain the local coherence of cells and at multi-cellular level to maintain the local coherence and differentiation of organs while at the same time organ interrelations permit the organism as a whole to show complex organisation, e.g. to permit the significant higher order redundancy characterising the co-ordination of lung, blood, muscular, nervous and hormonal systems that permit fight, flight and rest. Second, while high order redundancy may be accompanied by low order redundancy, as the case of living systems shows, it need not be. A computer CPU has much less low order redundancy than a memory chip but much greater higher level redundancy, as a casual glance at pictures of chips will show. Complex chaotic (and nearly chaotic) conservative systems, e.g. a steel ball pendulum swung over a pair of magnets under frictionless conditions, typically show relatively little low order redundancy, but a significant amount of high order redundancy. However, unlike the self-similar patterns associated with chaos, the local and global redundancies of living systems are typically highly non-self-similar across orders.[58]

We will also differentiate between local and global system redundancies, e.g. between those of the body and that of one of its organs. While the body has organs which are essential to its organisation, it is also intrinsically globally organised, and recall that this last is an essential feature of all autonomous systems. Modular systems segregate their redundancies into those within and those between their modules, while integrated systems do not (see Section 8 below), but integration can also be "patchy", across diffusely defined regions of a system, varying over time, as does that synchrony among neurones which Crick takes to define consciousness [45]. It is ultimately an empirical matter just how local and global redundancies interrelate to lower and higher order redundancies in particular classes of systems, though in general higher order redundancies will also have the larger temporal or physical scale, or both. For a group of components that support interrelations of a maximum finite order, eg. DNA components, the only way to obtain higher order relations is to add components, but it is also possible to have localised redundancy of very high order, e.g. if a localised sub- system supports a strange attractor (cf. note 58).

To be organised requires redundancy. But real systems show various combinations of high and low order redundancy, local and global redundancy. This provides an internal richness to the notion of organisation. It also undermines any attempt to provide a simple univocal redundancy correlate of organisation. The foregoing discussion suggests, for example, that we distinguish between the quantity and quality of organisation in a system, since adding more components to a system without increasing the order of redundancy among their interrelations does not seem to produce any increase in quality of organisation, just in quantity of compnents organised. We might associate the quality of organisation with the highest order of redundancy the system exhibits and the quantity of organisation with the total number of redundancies present across all orders. But while the former intuition is well defined, there is before us no apparatus for quantifying organisation in this way, particularly not one with principled grounding in the underlying notion of physical information. However, even the notion of organisational quality is rendered ambiguous by variation along the global/local continuum and by scale dependence. To the extent a system is modularised, for example, we regard its organisation as distinct from that of its modules, especially if the redundancies within one of its modules is higher order than that between its modules, but what of a system that is but patchily redundant across space and time, as our brains are? We therefore rest content for the moment with the following: The higher order the global redundancies involved at the mesoscopic scale the more organised the system. Computers and creatures are more highly organised than mechanical machines and machines are more organised than perfect crystals.

A significantly organised system is not maximally complex, because of its redundancy (more internally ordered than a gas), but it is not maximally ordered either, because of its higher order correlations (less ordered than a crystal). This is the intermediate region we have previously established (Sections 2,5) for type 4, living systems and represents the organised complexity required for AAA systems. There is at least one promising tool for capturing this kind of organisation.

As the examples of the Shakespearean sentence, PhD word, and programmed computer show, the higher order redundancy of systems is often, perhaps typically, hidden or buried, in the sense that it is not evident from inspecting small parts of the system or local segments of the dynamic trajectory of the system. However it can (in principle) be seen in the overall structure of the system, and/or in the statistics of its trajectory. For example, it is impossible to specify what a computer drawing a complicated picture is doing from examining either a few transistor states or a small section of the drawn lines; equally, the trajectory of a chaotic system is locally chaotic, but it is often confined to spatially restricted attractor basins. Because the information in such systems involves large numbers of components considered together without any possibility of simplification to logically additive combinations of subsystems (the systems are non-linear), computation of the surface form from the maximally compressed form (typically an equation) requires many individual steps, i.e. it has considerable *logical depth* [5, 93]. Of course, this measure applies whether or not we regard the order as epistemically hidden or buried. Formally, logical depth is a measure of the minimal computation time (in number of computational steps) required to compute an uncompressed string from its maximally compressed form.[59] Physically, the logical depth of a system places a lower limit on how quickly the system can form from disassembled resources.[60]

Bennett has proposed that logical depth, a measure of redundancy, is a suitable measure of the organisation in a system. However, while adding more components to a system at the same redundancy level will not increase the system organisation, only the size of the system organised, it will increase its depth because the sheer length of the sequence to be computed has increased. All sequences of n identical entries are intuitively equally trivial, however the depth of each string depends on the depth of n itself. This effect can be made negligible if we consider only relative depth: The depth of a sequence relative to the depth of the length of the sequence. The relative depth itself of a sequence of n identical entries is no more than the depth required to specify the entry itself (and negligible if the entry is 0 or 1). In the case of adding identical components to a system the relative depth does not increase since the depth of a compnent is already included in the original system relative depth. It is not transparent whether relative depth deals satisfactorily with all possible cases of this kind, but it is a reasonable, and plausibly sufficient, refinement of logical depth simplicitur to adopt.

Relative logical depth promises to capture the core of the intuitive sense of organisation we discussed above. More cautiously, relative logical depth seems strongly associated with the highest order mesoscopic global redundancy in a system. On the one hand, since low order redundancy can be produced with relatively simple programs (in the order 1 case, with a "repeat" command), high relative depth implies high order redundancy. On the other hand, from the examples above, it seems that high order redundancy requires considerable relative depth, since typical examples take a longer time to produce from simpler resources. Crystals are much easier to produce than are computers and, within the latter, otherwise comparable memory chips are much easier to produce than CPUs. However we do not know of a theorem that requires this. We will assume that higher redundancy order implies higher logical depth, other things being equal. Whether or not organisation requires anything else is somewhat unclear to us at present. It is, for example, unclear to us whether relative logical depth can be used to satisfactorily distinguish the kind of deep organisation exhibited by a computer running a complex coherent programme and that exhibited by a living cell, which is autonomous, defends an internal/external phase separation through a controlling boundary membrane, and is self- regenerating (and reproducing), all organisational features the computer lacks.

Lurking behind this issue are two important ambiguities in our inuitive concept of organisation; these need to be made explicit. The first arises from the fact that system redundancy orders are quite distinct from system levels. A cohesive system level is a dynamically grounded real constraint (structural or process) in a complex system which occurs when (and only when) cohesion emerges and operates to create organisation (note 6). A level acts as a macro filter, its formation requires a change in dynamical form, a phase change, and expresses a sharp shift in degrees of freedom (Section 2). Because a level may (and typically will) contain many components (e.g. many molecules bound into a structure, or circulating within Bénard cells), the same level may manifest or support many different orders of redundancy and the same order of redundancy may be manifested or supported at many different levels. (Since the use of "level" in the literature at large is multi-vocal and often undisciplined, and allows the hasty to reify talk of all kinds of levels, we shall confine the use of "level" to just the foregoing cases.) The first ambiguity concerns whether levels should play any intrinsic role in the concept of organisation. Consider two systems constituted of the same basic components and with equal measures of correlations among them, the first a single level system with subtlely correlated parts so that they mutually regulate and some control others, and the second a multi-level heirarchy with internally simple levels but cross-level correlations, none of which are control relations. If the notion of organisation is fundamentally concerned with just the extent of co-ordination in the sense of correlation then ex hypothesi these two systems would have the same measure of organisation. If, on the other hand, the notion of organisation also includes reference to the degree to which correlations are heirarchically organised by levels, as it will for many, then, the second system will be judged to possess the greater organisation. The second ambiguity in our concept of organisation concerns whether ordering of regulation or control should play any intrinsic role in the concept of organisation. If they are to then the first system will be judged to possess more organisation than the second. Here it will not suffice for systems to display correlations to exhibit organisation, they will need to display correlations distinctive of regulatory or control relationships. Finally, these two intuitions are often joined to emphasise the importance of heirarchical order of regulation or control to organisation. In this case the second system will be judged to possess more organisation than the first, since the first system can show only first order hierarchical regulation or control while the second system can show higher order hierarchical regulation and/or control. Our response is to recognise the usefulness of both kinds of ways of modifying the root concept, reserve organisation simplicitur for the correlations-only notion and introduce the terms "hierarchical organisation", "regulatory (control) organisation" and "hierarchical regulatory (control) organisation" for the others.

The importance of these distinctions goes beyond conceptual clarity. The distinctive features of the living cell introduced above all concern its heirarchical and regulatory/control character, which are dynamical features. This extends to characterising type 4 systems generally. But logical depth is concerned with organisation simplicitur, i.e. just with correlations; evidently it is silent about dynamics. If this is the case then organisation simplicitur will not suffice to pick out the features of type 4 systems which are crucial for understanding them, in particular for understanding their AAAR capacities. Plausibly these concern their distinctive heirarchical and/or regulatory/control organisation. To deal with these logical depth evidently needs to be supplemented with a dynamical account of depth, within the context of NPI. ("Evidently" because it is possible, though it seems unlikely, that the logical depth construction will prove to contain the resources for this task.[61])

How to dynamically ground logical depth is not presently entirely clear because it requires a way to physically quantify the notion of computational time, or, equivalently, of a computational step, and how to properly do this is not clear (note 60). However, when we do observe organisation we can reasonably infer that it is the result of a dynamical process that can produce depth. The most likely source of the complex connections in an organised system is an historically long dynamical process. Bennett recognised this in the following conjecture ([5] quoted in [93]):

> A structure is deep, if it is superficially random but subtly redundant, in other words, if almost all its algorithmic probability is contributed by slow-running programs. ... A priori the most probable explanation of "organized

information" such as the sequence of bases in a naturally occurring DNA molecule is that it is the product of an extremely long biological process.

However we should also note that higher order redundancy could arise accidentally as an epiphenomenon (a mere correlation), but then it would not be based on a cohesive structure and so its emergence cannot be controlled and it will not persist. It is also possible that cohesively based higher order redundancy could form accidentally via dissipation, e.g. as eddies in an expanding gas. Some of our own cosmic structure seems of this kind, gravitation providing the cohesive force. Thus, though we tentatively adopt (relative) logical depth as an important measure of system organisation simplicitur, we regard it as an incomplete account of the important system features we wish to characterise and look towards grounding it in NPI.

For the sake of clarity we pause here to note that the converse of Bennett's claim is not generally true: a system's being the product of an extremely long biological process does not ensure that it will contain a lot of organised information; in many environments the best adapted creatures might be relatively simple. Too much organisation can be as much of a disadvantage as too little, since organisation not only enables, but also restricts system variability by placing constraints on the system, making the system less adaptable to conditions the system is not organised to deal with. There appears, for example, to be an optimal range of system interconnectedness at between three and six connections per system node [134]. Furthermore, substantial portions of the information concerning the complex evolutionary history of the whole environment-system supersystem may be lost through species extinctions and not appear in its current physical information. And many of the simpler creatures may be long term survivors, like amoebae, rather than the recent products of selection. It is presently unclear whether or not our planet's evolutionary process is increasing the mean depth of physical information in its creatures, though this not occuring is compatible with depth increasing in some lineages (e.g. the mammals, [106]) and with increase in mean depth being the general evolutionary tendency but for external disruption, geological (volcano eruptions, etc.) and cosmic (comet collisions, etc.).

However, it is also important not to judge this issue by considering structures in isolation. It is tempting to conclude, for example, that many viruses are recent products of selection and yet simple. But viruses are not autonomous systems; they do not control themselves but need the higher order organisation of their hosts to reproduce. They can be simple because the control resides elsewhere. By contrast a specialised autonomous species may need its specialised food source for exergy and materials but may not need the food's organisation to control its self-regeneration processes. In fact, when food organisation enters autonomous organisation and is viable as such — as when a drug or virus is ingested — it often causes trouble precisely because the system does not then fully control the subsequent processes, although humans, for example, are dependent on their food containing some amino acids which they cannot synthesise themselves and to that extent have surrendered control of self-regeneration to their environment. The perhaps counter-intuitive but important lesson is that the order and organisation of the virus cannot be characterised for the material virus object alone but only for the entire virus-host system, even if the former can be separately structurally characterised, whereas the order and organisation of systems can be separately characterised, no matter how dependent they are on their environment, to the extent they are autonomous. (I.e. parasitism implies much stronger informational requirements than mere dependence.) More generally, for all sub-system/system relations, only in the respect in which a sub-system is autonomous can its organisation be considered separately. Internal organs/parts can range from autonomous sub-systems to complete parasites, in each case more or less co-operative; the exact relationship in each case is an empirical issue.

We can now explain Bennett's conjecture in more common biological terms. *Entrenchment* is physically embodied depth *per se*, with no direct implications concerning the historical origins of the depth. *Canalisation*, on the other hand is entrenchment resulting from a deep historical process (and also describes the process). Bennett's conjecture is, then, that cases of entrenchment are, most likely, cases of canalisation. This is an empirical claim.

Depth permits complex, organised control, e.g. as with programmed computers and DNA control of development, even in complex conservative systems.[62] There is still a problem, however, of how the necessary organisation might originate.[63] Since organisation requires both order and complexity, which "pull" in opposite directions, solving this problem is tantamount to explaining how the two can come into existence at the same time. Doing so involves introducing dissipative systems.

7. Dissipative Systems and Self-Organisation

Dissipative systems are open, non-equilibrium dynamically irreversible systems. Thus, as noted earlier, if they are to maintain their order, a fortiori if they are to self-organise, they need to export their entropically degraded (disordered) energy and import ordered energy as intropy and/or enformation. Entropy production is analogous to friction (friction is actually a special case) and systems tend to evolve along trajectories that follow the local path of least resistance determined by their constraints (minimise "friction"), since these trajectories have the lowest accessible energy levels. This leads to a minimization of the local rate of entropy production, commonly denoted by φ.[64] Under the right conditions, φ will be minimised if the system changes its state so as to produce macroscopic correlations. The classic exemplar is Bénard cell convection, in which convection in rotating cells replaces conduction along a temperature gradient (e.g. convection cells replacing vertical heat conduction in a slowly heated saucepan); here convection starts when φ is greater for the non-convecting state than for the convecting state.[65] It is a self-organising process, the spontaneous production of macroscopic higher order redundancy. Physically, this implies creation of new non-local spatial and temporal correlations among individual component (particle) trajectories. This is exactly what we observe in the transition from conduction to convection for molecular motions both within and between rotating cells.

The most commonly studied systems are near-to-equilibrium systems where departures from equilibrium can be linearly approximated [17]. In these systems the local statistical fluctuations are larger than or equal to the local gradient of intensive state variables, so they can for the most part be treated like equilibrium systems. The minimization of entropy production is global and locally isotropic.[66] The isotropic nature of φ in near-to-equilibrium systems severely constrains their complexity. By and large, their structure is under the control of their boundary conditions, e.g. Bénard cells and river standing waves formed with a fixed stony bed and bank shape, because the isotropy of φ implies that the system cannot act locally on itself to produce differential internal structure [96, 98]. And this also removes any historicalness from their constraints. This constraint on complexity is also a constraint on depth, and hence organisation. However, near-to-equilibrium dissipative systems are somewhat autonomous, since their internal order results from internal processes, and the order is relatively robust in the face of chance perturbations within or to the system. But their autonomy is limited by their inability to act on themselves, which requires further internal phase separations. This is reflected in analyses of near-to-equilibrium dissipative structures, which typically consider only the microscopic dynamics and the global dynamics, constrained only by the requirement of consistency, specified in terms of boundary conditions (see note 65).

In far-from-equilibrium systems, the situation is more complex, since local changes in the dynamics of the system can occur faster than the system can dissipate them or adapt to them. This creates the conditions for complex non-linear interactions to play a significant role within the dynamics of the system. However, the system still tries to follow the path of least resistance locally by minimising φ, insofar as it can. Prigogine [113, 114, 105] has conjectured that a far from equilibrium system will respond in such a way that it minimises the component of φ in the generalised direction of any generalised force applied to the system. Often, this will involve a change in the macroscopic state of the system that produces new information within the system. Other components of φ are not necessarily reduced, let alone minimised. The resulting potential can lead to further effects on other aspects of the system. And as long as far-from-equilibrium conditions are maintained, perturbation and re-organisation, even disruption on occasion, will tend to propagate through the system. This is quite different from near-to-equilibrium systems, in which all components of φ are reduced almost simultaneously, damping out any propagation of disruptions, and leading to a stationary state rather quickly. Far from equilibrium conditions, then, have more self-organisational possibilities, and consequently show constraint historicalness, i.e. their macroscopic constraints, the laws under which they operate, evolve/develop irreversibly over time.[67] This historicity is also a central characteristic feature of living, especially of learning, systems. And indeed far-from-equilibrium systems lead to much more diversity and complexity of response to perturbations (stimuli), and are much more conducive to the development of autonomous anticipatory systems, than either conservative or near-to-equilibrium conditions.

From an informational point of view, self-organisation through dissipation (Section 2) is the formation of enformation from intropy. Systems that are already strongly constrained (like nearly conservative or nearly equilibrium systems) have little variability to support this sort of transformation. Far-from-equilibrium dissipative systems, on the other hand, can produce increased enformation and depth through the complex and diverse but organised ways they can respond to external forces. As long as they are kept well away from equilibrium, the complexity and organisation of such systems can increase indefinitely. This provides the possibility of their acquiring the informational complexity required for autonomy, anticipation and adaptiveness (AAAness).

But we can go a small step further than mere possibility. A system would be more autonomous if it could resist external forces by itself repairing their deleterious effects or otherwise re-organising itself internally to make itself less vulnerable to disruptive forces. It would be even more autonomous if it could anticipate disrupting forces and prepare itself to counter them before they arrive. Something like this, in a generalised, rudimentary form, already characterises dissipative self-organising systems. The processes underlying self-organisation tend to reduce the entropy gradient that impinges on the system. Such a reduction would decrease the intensity of the entropy potential and consequent disruptive effects of both the gradient itself and of internal fluctuations [113, 114, 105, 122, 123]. Brooks has called this tendency of a dissipative system to re-organise itself to resist disruptive forces the *Principle of Compensatory Change* [13]. Although the ability of different systems to compensate varies according to boundary conditions and internal structure, all dissipative systems have a rudimentary ability to resist disruptions inherent in their dynamics. This ability is rudimentarily anticipatory at least inasmuch as it applies to generic disruptions, in particular to counteracting the effects of dissipation (see above).

On the other hand, it is a spontaneous response to disruption, enabling self-regulation, but able to sustain control processes only insamuch as it is constrained to occur in predictable ways that further support autonomy. More discerning anticipation requires further elaboration and articulation of this basic dynamical principle, first negatively toward avoiding damaging environmental circumstances and then positively toward those environmental circumstances where resources may be obtained for preserving developmental processes, thus leading to adaptive, anticipative systems. Once established, spontaneous compensatory responses can become embedded in the dynamics of the system through selection (either natural or artificial) or other entrenching processes, ready to be triggered by appropriate environmental input or their vicariant representations (see Section 9 below). In principle, compensatory responses could be designed rather than evolved, but this is an unlikely history for all but the most advanced cognitive cases. For example, Piagetian accomodation could follow either path, however only the spontaneous origin account seems likely to underlie the earlier stages of cognitive development. Similarly, in evolution it would be much more efficient to select preexisting spontaneous compensatory responses than to construct such responses piecemeal.

Living system complex organisation, and that in AAAR systems in particular, then, requires far-from-equilibrium conditions in a many-component interactive, non-linear system with structure and process dynamically emergent. Far-from-equilibrium dissipative conditions are both necessary for, and likely to produce and support, the degree of organisation required for AAAR systems and their characteristic self-organising processes. But this is not an equivalence since it is not an inevitable result of these conditions, as shown by the range of complexly organised systems lacking these properties. We have no neat formula for what additional constraints to add to single out the AAAR systems (or even just living systems) which, like that of dissipativeness, would be more fundamental than specifying their organisational character directly.[68] In any event, we next turn to briefly investigate one further relevant type of constraint.

8. Modularity and Organisation

The constraints characterising modular systems are such that the system dynamics can be expressed as an interactive product, the dynamical product of its intra-modular dynamics and its inter-modular dynamics. This provides a principled distinction between components of the system. We distinguish two kinds of modularity, based on two dimensions to cohesion relations, horizontal and vertical modularity (respectively *Hmodularity* and *Vmodularity*). Hmodularity obtains when there is a principled division of a system into contemporaneous spatial parts with enough unity or cohesion to each part for the system dynamics to be expressible as the product of the individual module dynamics and their interactions. This is how we currently design and model buildings and machines of all kinds (from homes to hotels, typewriters to television sets) and how we usually attempt to model both biological populations (the modules being the phenotypes) and their individual members (the modules being internal organs). Vmodularity, in contradistinction, obtains when a system's dynamics may be decomposed into the interactive product of its dynamics at different system levels. This requires the presence of relatively macroscopic constraints in systems sufficiently cohesive to impose extractable constraints on their dynamical behaviour. A crystal lattice, for example, resists disruption by thermal agitation of its lattice molecules and thereby constrains them to lattice interrelations, an internal organ has a similar relation to its member cells. Each instance of Hmodularity also involves many instances of Vmodularity, since each Hmodule must be sufficiently cohesive to be a dynamically separable part and so must exhibit Vmodularity with respect to its members, but Vmodularity may be system-wide (as with crystals, typically) and involve no Hmodularity.

Modularity means that a system satisfies certain constraints as a matter of internal cohesive dynamics without the need to invest further resources to ensure that the relevant correlations are maintained. Moreover it allows system resources to be used in different ways with less chance of interference or confusion, and it allows modules to be modified without modifying the whole system. Again, both of these features reduce the need for additional resources to maintain needed correlational structures. In particular, the regulatory resources required to maintain a high order redundancy, e.g. a system-wide operational coherency condition, are great and the risk of error correspondingly large; if instead an appropriate cohesive structure were present which dynamically constrained the system to satisfying the condition then the system redundancy condition would instead be nomicly assured and system resources could be invested elsewhere. For example, direct parallel connection is a simple and effective, and typically more reliable, method of maintaing output phase coherency of electrical generators than is sampling, computing discrepancies and corrective signals, and feeding them back to the individual generators, and similarly for direct mechanical governors; much of the cohesive structure of living bodies plays similar roles (e.g. cell wall cohesion as regulator of sodium- potassium balance). So we can expect to find modularity intimately connected with the expression of higher order redundancy, i.e. with logical depth. (However, we also note that, while modularity reduces resource demands, this benefit may in various circumstances also be outweighed by its costs, e.g. where there is selective advantage to altering modularity.)

But while Vmodularity makes it possible (though not necessary) for a system to display global logical depth while remaining locally shallower or more simply organised, Hmodularity makes it possible to trade-off greater organisational depth in individual modules against shallower organisation in the interrelations among the modules. H and V modular systems combine these two opposite-leaning characteristics in various ways to make systems that can display both deep modular subsystems and deep global organisation. Living systems are generally (if not always) of this sort.

For completeness we should recognise process modularities as well as the spatio-temporally based H and V modularities. This is especially so for autonomous systems, for whom its process closures play the key role. As for modularity simplicitur, process modularity occurs when a system's process dynamics can be expressed as a product of its intra-modular and inter-modular process dynamics, and again we may have both H and V process modularity following analogous constraint definitions. A serial computer completing a programme with nested iteration loops must exhibit V process modularity while a parallel distributed processing machine must exhibit H process modularity (though in both cases the other modularity may also be present). With our currently severely limited capacity to model the dynamics of complexly organised systems, it is often more useful to directly model system processes and then process modularities typically provide important insights into system capacities. Modularities of communication and decision making within a business organisation, e.g., reveal ways that its members do not interact which may provide important insights into both its capacities and limitations. (For a particularly important class of cases see the ultimate paragraph of this section.) But, so far as we can see, all H process modularity arises from and requires appropriate Hmodularity and similarly for V process modularity. So we proceed to discuss H and V modularity, treating process modularity as derivative on these, while recognising the importance of process dynamics to autonomy and so to all AAA systems.

Modular systems have internal cohesive structure that constrains system dynamics. These constraining cohesions are required for both regulation and control relationships within the system, which themselves impose further variable constraints on system dynamics. Vmodularity results in system levels which in general produce vertical control asymmetry. A higher level constrains lower level dynamics (as a crystal lattice constrains the behaviour of its atomic constituents), will often regulate it through feedback (e.g. coherence of crystal vibrations) and sometimes it will also control the lower levels in important respects (top-down control, e.g. brain control of muscle). But it will also typically be true that lower level dynamics will constrain higher levels (as electron orbital dynamics constrains crystal angles), may regulate them through feedback (e.g. catalysis of chemical reactions) and might control certain aspects of the higher level (bottom-up control, e.g. indirect control of volume Hebbian learning through local NO release). And for many properties it might be the case that there is no control asymmetry involved, simply mutual constraint through interaction, e.g. of oscillatory behaviour in a system of small oscillating springs connected to a common rigid, but moveable, bar. This latter interaction-only condition will be the common case among Hmodular components at the same level, e.g. among cells of the same organ. Whenever there is a control asymmetry we shall speak of hierarchy relationships, with the direction of the hierarchy being the direction of control asymmetry (cf. [52, 118]; Dyke's usage is confined to the rare special case where constraint is one-way only, [49]). Commonly among living and human engineered systems a heirarchy is specified by assembling cohesive combinations of Hmodular components, e.g. building organs from cells and bodies from organs. But it is possible to have at least dynamically distributed feedback regulation that does not require Hmodularity, e.g. the distributed rate dependency phase separation of Belusov-Zabotinsky chemical reaction structures.

Some explanation for the modularity features of living systems is in order. One possible explanation is that the dynamics of self-organisation itself leads inevitably to an H and V modular form of organisation. Whenever new enformation is produced through self-organising processes, that enformation is distributed in the form of large scale temporal and spatial correlations within the system. This amounts to a new level of organisation that did not exist in the system before local activities became globally co-ordinated. This does not imply that the system organisation must be modular; e.g. a system consisting of a single Bénard cell will show global organisation, but has no modules, H or V. Many other classic dissipative structures show no modularity. Nonetheless, many do, simply as a result of their dynamics. Typically, as the system self-organises, it fragments into cells, which in turn self-organise, and so on. In the case of Bénard cell formation, e.g., an arrangement of many smaller cells provides far more efficient heat convection (and so dissipation) than does one large cell and so as the Bénard system is forced more strongly it forms cells of dynamical necessity. Another possible explanation for modularity is that the origins of living systems involved the separate evolution of different modules that gained a competitive advantage by organising together. This theme is now common in origin of life scenarios [51, 138], and is also a popular explanation of some aspects of the structure of eukaryotic cells. Holland locates the reason for this character principally in the power of cross-over and transposition in genetic algorithms [67]. This account places the explanation in historical circumstances allied with genetic constraints. It explains why cooperation was superior to development of the same capacities independently because of the advantages of distributing adaptation among trait-producing modules rather than waiting for useful gene combinations to emerge with an otherwise disorganised whole. A further possible explanation looks to the adaptive advantages of distributing work among modules rather than exerting central control over otherwise disorganised parts. As noted above, the advantages probably derive from resources saved or freed up for other uses. (Cherniak [20], for example, provides a nice argument that memory storage should be somewhat, but not too, modular for efficient retrieval.) There may be other advantages as well. In sum, because of the systematically interrelated possible states modularity economically provides, modularity is essential to the kind of co-ordinated organisation and capacities possessed by AAAR systems but most likely dynamical, historical and adaptive considerations are all involved in explaining the prevalence, and particularities, of modular organisation.

Nonetheless, it is importantly also typical of living systems that their modularities, both H and V, are approximate only, both through interpenetration of cohesive structures (e.g. the interpenetration of the hormonal and cardio-vascular sub-systems) but more importantly through cohesive structures being modifiable, and modified, as a consequence of system dynamics. The operation of the nervous system, for example, profoundly affects that of the hormonal system, and vice versa, with the affects not just confined to their states within fixed neural and hormonal dynamics but includes these dynamics themselves (e.g. through new neural connections and changed glandular production relations). This is where a process dynamics specification may be of particular use, since process organisation may be preserved across such changes — indeed, it may be that the changes occur precisely to ensure that (as when the body alters its local physiology to support a higher order process, in switching to fat consumption under performance stress, or stimulating new dendritic branching to satisfy reward demands). Sometimes modularity shifts spectacularly, as in the emergence of a new macroscopic, internally specialised form from aggregation of individual identical cells in the slime mould

Dictyostelium under starvation conditions ([55, 65]; this is also an example of preserving higher order regenerative process). To reduce complexity and increase ease of regulation or control, currently we largely design such plasticity/flexibility out of engineered systems in favour of rigid modularities.

9. Adaptation and Anticipation

We conclude our discussion by bringing adaptation into the dynamical information framework we have established. Construed generally, the process of adaptation is a system-environment open-loop interaction that yields system modification such that system unity is at least preserved and system autonomy and internal system information increases. We shall say that a system with a capacity for that process is adaptive in that respect and call the resulting system modification an adaptation. Recalling our discussion of Section 2, adaptations may involve either system reorganisation, e.g. those that result from assimilation of new perceptual information, or system self-organisation, e.g. those resulting from perceptual accommodation or appropriate genetic change. From this principled perspective we may consider all these generalised "learning" processes. Indeed, even if we restrict this term to self-organisation it will comprise both genetic algorithm processes and various psychological processes. Various versions of a unified conception of evolution and ontogenesis of this sort have been claimed or suggested [9, 50, 70, 72, 109, 110, 111] though the specific roles of all of the re- and self-organisation processes is ultimately an empirical matter. Since an adaptation need not be deep or especially complex, though it often is, adaptation varies somewhat independently of system complexity, order and organisational depth. Adaptation introduces the first semantic notions into information theory, in the form of significance, so that something more immediate can be interpreted as a sign of something else [8, 86]. Typically, for example, adaptive behaviour is not directed immediately at self-preservation, but only through some means whose significance is the increased likelihood of self-preservation. For example, the adaptive significance of food seeking behaviour is that food is required for life and biological reproduction.[69]

In evolutionary theory a trait is an adaptation if the process producing it involves essential reference to what having the trait achieves (in terms of preservation of the lineage). This may be a narrower conception than that above and, while it may be seen as appropriate for the AAAR systems involved, it is unclear how to specify it in a principled manner. (Rather, it invites etiological talk of the purposes of capacities and functions, which quickly becomes a morass; [21].) Any trait that is produced by natural selection is an adaptation. The converse is more controversial. Orthodox biological theory asserts that all or nearly all adaptation results from natural selection, to the point that an adapted trait is not considered an adaptation unless it has been selected. If the trait is merely there and happens to aid survival, it is just that, and not an adaptation; it is said to be adapted (or, confusingly, adaptive). However, this restriction is questionable and impossible to specify non-arbitrarily; it would again be better replaced by a principled distinction, such as that above. However, adaptation as an evolutionary process in a lineage is traditionally distinguished from adaptation as developmental, physiological and behavioural processes in an individual organism [16, 137]. But as we have shown, the two senses can be unified by subsuming both processes under a general principled concept of enhanced information and autonomy which perform the same ultimate function for lineages as individual adaptation performs for organisms. Indeed, from this perspective biological reproductive capacity and adaptiveness, 'AR' of AAAR systems, are necessary (and plausibly sufficient) conditions for lineage autonomy.[70]

Adaptation implies a correlation between what is adapted and its environment, which in turn implies mutual information with the environment, called the *information of adaptation* [39, 42, 136]. A measure of the complexity of adaptation of a system, this information places certain constraints on the system-environment super-system that enhance system survivability by making strategies that lead to destruction less likely, and strategies that lead to self preservation more likely. The information of adaptation is a mutual constraint between the system and its environment that shapes the *fitness landscape* for the system's survival. The solution of problems posed by the environment requires the production of the information in these constraints. This can come about either through chance adaptation, or through a more inventive anticipatory process. The latter enhances adaptability, but it requires greater informational resources, both in terms of complexity and depth (organisation).

While in genetic evolution environmental information constrains the genetic information that survives by acting as a filter, panselectionist evolution is highly idealised; in fact not all useless traits are eliminated, and the most adapted traits do not necessarily survive. Evolution can go faster than genetic equilibrium is reached, allowing non-optimal traits to survive, as is suggested by the wide phenotypic diversity of most species. This latter can be an advantage where adaptability is concerned [39, 40], but is to be contrasted with the achievement of behavioural adaptability through endogenously directed re- and self-organisation [63, 26, 28, 30]. If evolution tracks the optimal fit between lineage and environment, then the lineage stays in genetic equilibrium with the environment, and there is little or no genetic variability to allow for changes in adaptive strategy. (Although an optimally adapted species could dynamically alter its environment so that a local fitness maximum became instead a saddlepoint, permitting further fitness ascent.) On the other hand, if evolution is not optimal it is not entirely environmentally driven, and possibilities arise for increasing the information of adaptation through self-organisation in adaptive space. This might happen, for example, through combinations of prior adaptations [44], or through the adaptive martialling of pleiotropies [2], though neither Conrad's nor Baatz' formulation distinguishes non-optimal evolution from the previous option. The extent to which non-optimal adaptation occurs is an empirical issue that requires quantification of the relative strengths of selective and other self-organising forces to resolve.[71]

Ontogenetic (developmental-psychological) adaptation is similar to evolutionary adaptation in requiring increased mutual information with the environment and again re-organising and self-organising processes are able to produce new information to be selected. In the case of psychological adaptation selection on vicariant representations of possible strategies in expected environments [9, 19] plays a central role and provides it with distinctively powerful articulation and usefulness, with self-directed vicariant learning playing a particularly striking role [29, 30]. Perhaps the bulk of the adaptiveness of thinking organisms over their unthinking competitors resides in this capacity.[72]

While adaptation does not require increased complexity or organisation, it produces both through the production of the information of adaptation. Conversely, the main advantage of complex organisation comes from its contributions to *adaptability*. Adaptability is second and higher order adaptation, the capacity to adapt adaptations. It is advantageous in variegated environments (both spatially and temporally) since no one adaptation can be but temporarily effective. Adaptability is inextricably linked to information processing ability [43] and through that it is intimately linked to intelligence ([71] and note 73 references). Whereas adaption tends to constrain expression of traits and behaviours to those that are functional in a given environment, through incorporating the right kinds of higher order organisational constraints (see below), adaptability permits the selection of traits and behaviours with both greater variability and greater specificity, thereby achieving the advantages of both non-optimal and optimal adaptation.

Adaptability has been operationally defined [43] as the uncertainty of the most uncertain environment it can tolerate. However this fails to distinguish between genuine adaptability and merely increasing system stability by closing off the system as much as possible, rendering it, stone-like, impervious to interaction. Genuine adaptability requires that toleration should be understood in terms of autonomy: system regenerative self-maintenance at a level sufficent for system survival. (This depends on the identity conditions of the system.) Adaptability is then enhanced by increasing the uncertainty of the most uncertain tolerable environment, which requires building features into the regulatory design of the system corresponding to as wide-ranging dynamical patterns shared by the class of tolerable environments as possible and using these as the basis for anticipative adaptations in response to local, short-term perceptual information. While the latter represents an immediate, short-term increase in the information of adaptation typically concerned with organisationally shallow features, the former represents a crucial long-term extraction of mutual information which corresponds to much more deeply organised environmental features. This kind of information requires in turn complex organisation within the system to express it. So we associate increases in adaptability with increases in organisational depth, with the greatest adaptability expressed in cognition. Moreover, responsiveness on this basis is clearly strongly anticipative, the system uses its higher order invariant patterns and current information to anticipate the current course of events and so adapt accordingly (typically, but by no means exclusively, behaviourally), with the strongest anticipativeness found in cognition. Finally, self-directedly organised adaptability of this kind provides the foundations for system intentionality, and cognition.[73]

10. Conclusion

Dynamical systems show a bewildering variety, many of them with highly complicated characters and capacities. We have aimed to bring some principled order into this variety, and to provide a principled basis for the analysis of dynamical system capacities. In particular, we have focused on developing a principled basis in dynamical systems for understanding the capacities of autonomy, adaptiveness and anticipation which, as we understand it, are the root capacities for life and intelligent systems.

Autonomous systems show all the hallmarks of life and intelligence. The essence of the living cell is its ability to actively maintain a phase separation between its interior and exterior of a kind which permits compensatory regulation over what passes through its boundary and in a way that permits self-regenerating control of internal production and biological reproduction, including this boundary-maintaining regulatory capacity itself — i.e. its essence is its autonomy. This basic autonomy is already highly organisationally complex and intrinsically anticipative in its compensatory dynamics, which also provides it with its adaptive capacities. Moreover, because of its reproductive capacity it belongs to a lineage whose autonomy derives from cellular adaptiveness and self-reproductiveness. Similarly, at a higher degree of organised complexity, intelligent systems display complexly organised internal control of anticipative response, conditionalising it on many subtle signals and, to the degree their control is itself adaptable, they are able to modify it and thus learn, with the most deeply organised systems capable of self-directing learning (learning how to learn). Their responses are aimed at solving a problem normative for them: How to act so as to maintain or enhance their autonomy, for that is a sine qua non of their continuing identity as that kind of organised system. If we assume that they have at least one way to evaluate their success or failure in this, that is, at least one reward signal (e.g. pain/pleasure), then they already display an epistemic relation to their initiating signals, taking a signal to signify the appropriateness of their anticipative response to it (downstream modulation) for this end, about which they could be in error. (This is where lineages part company from phenotypes as individuals.) In this manner they generate for themselves a semantic content for a signal, namely what is thus signified. (Epistemicly their end is to maintain or enhance reward, while its dynamical ground is to maintain or enhance their autonomy.) And, combining self-directing adaptability with epistemic semantics, they display basic intentionality. (Note that on this account contemporary computers are not intelligent or intentional systems at all because they lack autonomy, having no self-significant, epistemic capacities or functioning of their own but rather being entirely derivative from our attributions of content to their states in virtue of our uses of their formal functions; because of their grossly simplified dynamics they are misleading models of what it is to be intelligent or possess a mind.) This provides, we believe, the principled basis for a realistic and powerful conception of intelligence in general, with human intelligence properly seen as a partial special case.

Notes

1. This approach is not novel with us, but we claim a greater precision and integration than is found in previous work. The idea of bringing together dynamics and information theory has roots in discussions of Maxwell's demon ([92, 37] for references), and has been developed by Brillouin [11], Landauer [87, 88], Bennett [3, 4, 5, 6], Gatlin [56], Layzer [90, 91], and Küppers [86]. Kampis [79] has applied the idea of self-organisation to the same set of problems we deal with. Recent papers by Kampis and other representatives of the general approach called internalism, or endophysics, have been collected by Matsuno [97]; while this approach parallels ours in several ways, we do not adopt the subjectivism of some of its representatives, see note 27.

2. In their technical dynamical usages, constraint, regulation and control are distinguished as follows: Constraint refers to any dynamically based restriction on system access to regions of its state space, regulation to constraint satisfaction through dynamical feedback relations (as opposed to simply eventuating from the stability of the unperturbed dynamics), and control to regulation obtained by a process of comparing relevant system state conditions with a reference condition, and generating a correcting feedback signal accordingly. We do not pursue these important distinctions further here, nor the subtleties of deciding which obtains (as opposed to system behaviour *as if* it obtained) because, while of great importance for understanding the precise ways various kinds of adaptiveness are realised [23, 28, 30], they are not crucial for the more general foundational account we are developing here. Instead, we will sometimes refer to "regulation and control" where we wish to avoid the distraction of a more detailed analysis, but more often utilise a more colloquial, vaguer sense to the term "control" under which it means (roughly) "sufficiently constrained" — as in "a computer is controlled by its programme" — where the precise means of constraint can encompass any or all three of constraint simpliciter, regulation and control as defined technically, and leave context to sufficiently disambiguate what is involved.

However, these terms can also be used in a more a-dynamical functional sense where "regulation" means only "yields (roughly) constant output for varying input" (where the "roughly" encompasses the usual considerations of allowable transients and nearness to constancy to count as sufficiently stabilised) and "control" means "regulated in a reference-conditioned way". Used thus they are multiply dynamically implementable and so imply nothing about the specifics of their dynamical implementation from application to application. The tendency to exploit, but not recognise, this ambiguity of functional/dynamical sense, and the failure to properly tackle the relations among dynamical and functional specifications, has led to many unnecessary difficulties in discussing complex organised systems ([68], [62, Sec.IV.2.3]). Setting aside the technical mathematical use of the term, here we eschew functional talk in place of talk of interactive capacities, which are their dynamical underpinnings, unless the context makes it clear either that there this is the colloquial meaning of "function" (as in "physiological functioning") or that we are explicitly referring to input-output characterisation.

3. Descartes' approach to the problem was so crude that it is a caricature. Supposing no conceptual connection between mental and physical activity, Descartes relegated them to separate substances. Physical activity was constrained by various fluid pressures working through tubes connecting parts of the body. Control was organised in the mind, and exerted through pressures the mind placed on the pineal gland. This control was then distributed through the body through the tubes. (For Freud's version see [112].) Replacing "nerves" for "tubes", suitable electrochemical activity for fluid pressure, and computation for mental activity, Descartes' position is not so far from contemporary cognitive science (except that the shadowy notion of essentially separate substances has been dropped in favour of property essentialism to justify the dualism). The theme is common to a variety of contemporary idealist approaches to mind and language ([95, 26] — where it is labelled "internalist").

4. See e.g. [10, 15, 31, 32, 73, 72]

5. The dynamical approach to control theory is not without its own problems. It requires considerable ingenuity (and/or luck and patience) to implement a general dynamical approach to specific process problems. In particular, dynamical solutions to different problems cannot necessarily be added together to solve a more complex problem that is the logical sum of the two problems, because of interfering interactions among the dynamical parts of the system. In the Cartesian model, a solution to a functionally specified problem that is computationally solved is not "lost" when the problem is added together with another functional problem that has been computationally solved, as long as the solutions are not strictly inconsistent. For example, a Cartesian robot that has computational solutions to the problems "move towards a light source", and "avoid barriers" thereby has a computational solution to the corresponding joint problem. However a robot with individual dynamical solutions to each of the corresponding process organisational problems ("generate a smooth interaction flow that converges system location on a light source location" etc.) might not be able to solve the combined problem because its dynamical solutions to avoiding barriers and moving towards a light

source might require the robot to use the same resources in incompatible ways. There is a corresponding difficulty within Cartesian design, however: given a solution (algorithm), it is not always clear what problem the solution solves. Except in very controlled circumstances (equivalent to a very restricted problem), not every contingency can be rigorously accounted for, and a given algorithm might turn out to be a solution for a much more restricted, or even different, problem than expected. This difficulty is well-recognised by programmers, and has been paid homage to in the sardonic labelling of some program bugs as "features".

6. System cohesion refers to the dynamical stabilities arising from the constraints which a system of a particular non-linear dynamical kind imposes on its constituents. The essential idea of cohesion is that of an emergent system dynamical property that is insensitive to relevant local variations (e.g. thermal fluctuations) in the system components, including in those non-linear interactions that formed it [35]. For example, a kite, but not a framed soap bubble, has noticeable lift in a wind because the cohesion of its surface molecules successfully integrates the collisions with air molecules and transfers it to the frame. This is in turn different in kind from the (still dynamical) communicational interactions that constitutes the cohesion of a flock of birds, and from the non-cohesive but correlated wave pattern formed in a boat's wake. Note also that while the kite's cohesion is primarily expressed as a *structural* stability, that of the bird flock is expressed primarily as the stability of the flocking *process* through flight path changes. Living systems are primarily characterised in terms of their process organisation. Their structures may change, and must change somewhat whenever their adaptability is manifested; the more organised their adaptability the higher order the cohesive processes that characterise them (see below and [27]–[30]). In [23, 35], we have shown how cohesion grounds both system properties and individuals.

7. There have been a number of attempts to develop a characterisation of actively individuated systems related to the concept of autonomy outlined here, though there is considerable diversity in the details. Maturana and Varela [99] present a theory of *autopoeitic*, or closed self-reproducing, systems based on cells as paradigm examples. Also using cells as the primary model, Rosen [117] develops a mathematical theory of self-repairing systems he calls *metabolic-repair* systems. Bickhard [8] contrasts energy well and far-from-equilibrium systems, and labels far-from-equilibrium systems whose identity is process-based *self-maintenant* systems. These had been much earlier noted by Fong [53, appendix]. Ulanowicz [133] and Smithers [129], to our knowledge independently of each other, both speak of a class of *autonomous* systems described as *self-governing*. The conception of autonomy used here (and developed further in [24, 30, 23]) is most influenced by the work of Rosen, Ulanowicz and Bickhard, however much of the detail of the analysis is original. (See e.g. [22, 27] for a critique of autopoiesis as too focused on first order structure and Smithers as too confined to first order dynamics.) Fong, whose main work was c.1970, has an overall approach similar to ours in outline, e.g. in ground-

ing explanation of more complex systems in physical organisation and information, but lacks the idea of autonomy and other later technical notions we employ. (However he has developed a bracing breadth of vision that we hope someday to match.) The concept of autonomy also has a long-standing definition in mathematical systems theory which is closely related to our own, see note 10.

8. For an initial account see [26, 70], [27]–[30].

9. This is a technical sense of "information" used in information theory. It is really a measure of the capacity to carry meaningful information.

10. That is, in regenerating itself a system must also regenerate its capacity to reproduce, and in reproducing it must confer on its offspring the capacity to regenerate as well as reproduce. (Note that for many systems regeneration may include component reproduction, e.g. cellular reproduction to replace component cells in multicellular organisms.) AAAR invariance allows us to relate our notion of autonomy to that found in mathematical systems theory. There "autonomy" is given a formal meaning: A system S is autonomous if the functional dynamics of S does not contain time explicitly [59]. On the one hand this definition is wider than the meaning we have given to the term here because it includes systems which are merely passively stable, like a brick, as well as those which actively maintain themselves at a far-from-equilibrium dynamical stability, as all living systems do and as our AAAR systems do (see below). On the other hand, if the definition is read as requiring that every aspect of the dynamics of an autonomous system be time invariant then it will be narrower than our own because actively self-regenerating systems include adaptable systems and all adaptable systems have the capacity to modify (adapt) some parts of their dynamics to better suit their current situation, so that in a time-varying environment (the typical biological case) at least some of their dynamics will be a function of time. However, it follows from the text remark that there is a privileged specification of the dynamics of these systems which is time-invariant, namely that it is AAAR. This corresponds to the whole-system level of cohesion or scale for these systems. Furthermore the time-varying components will only correspond to autonomous sub-systems when these latter also satisfy such an invariant specification and hence exhibit appropriate cohesion at their whole-sub-system scale. Thus, excluding passive stability, and with the definition read as stating that a system is autonomous just in respect of its time-invariant dynamics, in the respects and at the system scales at which they occur, the two concepts coincide.

11. One way to do this, perhaps the only way, is to allow the non-programming dynamical plasticity of the material computer to actively modify programmes, i.e. to seriously embody the computer intelligence – see e.g. [7, 26].

12. It is sometimes possible to provide sufficient ordered energy in the specification of initial conditions that the system will proceed with the organising process as a

matter of moving toward equilibrium, e.g. by having the bodies in the latter case initially moving on collision courses at above-threshold energies, but this is just to replace one kind of ordered input by another.

13. We call symmetry-breaking through fixation of uncontrolled variations via internal system dynamics *spontaneous* for the reason that, although the order implied by the symmetry-breaking arises immediately out of the system dynamics, it is not wholly controlled by the system dynamics, much as a person's spontaneous outburst is not in itself deliberate, but nonetheless arises out of their underlying intentions and personality.

14. In the reorganising cases the cost to system organisation through unavoidable dissipative losses in these processes are typically assumed sufficiently small and ignored because they do not disrupt the reorganising process. They may occasion disruption elsewhere, when one part of a system is sacrificed to supply order to another, or they may be negligible or non-existent and the systems be quasi- or strictly conservative, as in reversible computations (see note 17 and Section 5 below). A careful exposition would also distinguish systems whose organising processes are equilibrium-forming, as in the crystal case, and those that occur in far-from-equilibrium dissipative systems, such as living and AAA systems (see below). The perfect crystal is a very simple organisation, its structure does not allow new possibilities of self-interaction without breaking cohesion (cf. note 66 and text), so there is no order increase (in the sense of [89]); both equilibrium and exothermic formation may be sufficient conditions for this organisational simplicity. Certainly, all living systems and complexly organised designed devices require endothermic formation, and the crystal quickly goes to an equilibrium state thus reducing intropy (system variability, see Section 4) almost immediately, suggesting that equilibrium, in contrast to far-from-equilibrium steady state, systems are degenerate.

There are no reversible organising processes proper since there is no increase in system order, hence there is no new organisation without dissipation. In the reversible case there may be shifts of order from one part of a system to another and these may mimic an organising of one part, but such processes can only ever have a misleadingly derivative status with respect to organising processes proper.

15. Note that there is no restriction on the size of the initiating fluctuations in relation to initial system scales, although fluctuations which are too small may be impotent and fluctuations which are too large may so disrupt the system that no ordered outcome eventuates. Typically, but not necessarily, the variation will be initially on some microscopic scale relative to system cohesion scale, which will explain why it is not controlled by the system. Water flowing over a stony river bed, for example, forms and re-forms standing waves as small flow fluctuations are amplified and fixed by the flow.

16. Thus in the terms of Section 4 below, self-organisation creates a cusp in the H_{MAX} curve rather than the incremental increase of re-organisation. At the same time H_{ACT} will also increase, though more slowly, providing for an increase in order and organisation.

17. This process is dissipative, though only described formally; in any physically realised version physical information must be lost to the system in the process. Similarly, any physically realised version of the first, standard net case would also be dissipative, although this is ignored in the standard description given (cf. note 14). If the learning process is reversible in the first case the order can only be transferred from, not copied from, the environment. Technically, copying is not necessarily irreversible because of the logical possibility of reversible computation [54], but it requires the production of waste memory, required to be kept in reserve to perform the reversal. When the memory is erased, the reversal becomes impossible to perform. In real cases, there is no significant memory. On the other hand, the possibility of reversible copying and then memory erasure that eliminates the connections required to make the original order shows that order transfer is possible through irreversible change as well.

18. Although the distinction between re- and self-organisation seems clear-cut, subtle issues arise in its application because of the internal complicatedness of living systems. Some of these have been initially explored in Collier [33]. Consider, e.g., a change which allows merely molecular DNA information to become genetic information by acquiring a control or expressive role (or, more variably, old genetic information to acquire a new control role), or which allows genes that are not phenotypically expressed for developmental reasons to become capable of phenotypic expression, their information thus being "promoted" from genetic information to phenotypic information. If these changes are analogous to the addition of new like molecules to the established crystal structure in the crystal-forming case they cannot be claimed to create a new cohesive level with new macro filtering, and hence not be self-organisation, but if they are analogous to the addition of impurities whose consequence is to change the crystal structure radically enough (say, introduce new modular regional sub-structures) then they can claim to be self-organisation. How much genetic and phenotypic change is enough to make good a claim for self-organisation? Though we believe that principled answers are always available for such questions — answers which illuminate precisely where cohesion operates, how autonomy is organised, and so on — each case must be carefully analysed before a decision can be reached.

Note also that while self-organisation creates new kinds of organisation in a system because of the level-formation involved, which particular kinds of organisational changes are created depends on the detailed dynamical nature of the system in which the self-organising process is occuring (how highly organised already, how energetic its bondings, how hot, etc.). Both crystals and embryos may self-organise, but the organisational results are very different.

19. This last characteristic has been called *downward causation* [18, 130]. See also note 24 and text. For an introduction to far-from-equilibrium dissipative systems see [113, 114, 105], and for self-organising systems see e.g. [115].

20. And moreover must have an autonomous organisational design. This organisation *could* be of the sort found in type 3 systems, i.e. effectively pre-programmed matching and selection using huge "lookup tables", so to speak; however for clear, if deep, reasons in our evolutionary world it is essentially absent short of human technological design and, as we remarked of computers above, it can only appear there in a very limited way. [7] reinforces our conception of AAA systems with their recent study of formalised versus real dynamical metabolisms. They emphasise that actual dynamical metabolisms are thermodynamic processes which cannot be fully captured in formal notions like process closure and that AAA dynamical possibilities, such as self-organisation, thereby elude formal capture. Strong artificial intelligence models are similar to the coin-sorting case, they only reorganise, not produce new (amounts of) organisation.

21. Cf. e.g. [69, 70] on Piaget. From this perspective it is an empirical question how distinct evolution and development processes are and some contemporary evolutionary theorists see evolution and development as only conceptually distinct. There are many real world cases that have elements of both, and in many of these self-organisation evidently plays an important role (note 71 and [13, 80, 119, 41]). Intellectual development need not be under either genetic or environmental control, nor need it be a simple sum of the two [30], something Piaget constantly emphasised [69] but is still a matter of controversy (e.g. [131]).

22. See also notes 34 and 41. Of course, which system aspects we choose to model on any given occasion is our choice — but the model's domain of empirical adequacy is not.

23. In the case of Bénard cell formation, e.g., we only have models which will predict the existence of the process because of the breakdown of stability conditions (see [41], for exposition) and [55], provides a detailed and insightful analysis of the comparable modelling problems arising in the case of slime mould aggregation. Herfel and Hooker [65] generalise this analysis, in particular seeing in it an explanation of why understanding scientific revolutions, modelled as self-organising commitment phase changes, are impossible to understand in terms of simple rational rules.

24. Seth Lloyd has compiled 31 ways to define complexity [74].

25. The label "computational complexity" is somewhat misleading since there are other notions of computational difficulty that might be called complexity, in particular the space, i.e. memory, and time resources, required for a computation. In any event, on the original definition

$$\text{length}\,(\sigma_s) \;=\; \min\{|p| \;:\; p \in \{0,1\}^* \;\&\; M(p) = \sigma_s\}$$

$$= \min\{|p| : p \in \{0,1\}^* \ \&\ f(p) = s\},$$

$|p|$ being the length of p, which is a binary string (i.e. $p \in \{0,1\}^*$, the set of all strings formed from the elements 1 and 0), and M being a specific Turing machine, and f being the *decoding function* to recover σ_s from p and then s from σ_s. This definition requires an $O(\log n)$ correction for a number of standard information theoretic functions. A newer definition, now standard, sets length(σ_s) to be the input of the shortest program to produce σ_s for a self-delimited reference universal Turing machine. This approach avoids $O(\log n)$ corrections in most cases, and also makes the relation between complexity and randomness more direct.

26. For a technical review of the logic of algorithmic complexity and related concepts, see [93]. The complexity of a program is itself a matter for algorithmic complexity theory. Since a universal Turing machine can duplicate each program on any other Turing machine, there is a partial recursive function f_0 for which the algorithmic complexity is less than or equal to the algorithmic complexity, plus a constant involving the computational overhead of duplicating the particular M, calculated using any other f. This is called the Invariance Theorem, a fundamental result of algorithmic complexity theory. Since there is a clear sense in which f_0 is optimal, the Invariance Theorem justifies ignoring the language dependence of length(σ_s) (but see previous footnote). String maps of highly complex structures can be computed, in general, with the same computational overhead as those of simple structures (the computational overhead is nearly constant), so for complex structures (large C_I) the negative component of informational complexity is negligible. Furthermore, in comparisons of algorithmic complexity, the overhead drops out except for a very small part required to make comparisons of complexity (even this drops out in comparisons of comparisons of complexity), so the relative algorithmic complexity is almost a direct measure of the relative informational complexity.

27. This is in contrast to its only being comparable via correspondence in the infinite limit, which is the only case in which the computational overhead, being a constant, is infinitesimal in proportion, [84, 93], and is therefore strictly negligible.

28. The converse is *not* true. Arbitrarily long substrings of non-computable strings (and, for that matter, incompressible finite strings) can be highly ordered, and therefore computable, but the location and length of these highly ordered substrings cannot be predicted from earlier or later elements in the string. In general, the incompressibility of a string does not imply the incompressibility of its substrings. Since it is possible to change an effectively random string into a compressible string with the change of one digit and yet, intuitively, the change of one digit should not affect whether a string is random, randomness of finite strings of length n is loosely defined as incompressibility within $O(\log n)$ [93, p.201]. By far the greatest proportion of strings are random and in the infinite case the set of

random strings has measure 1. It is also worth noting that there are infinite binary strings whose frequency of 1s in the long run is .5, even though the strings are compressible, e.g. an alternation of 1s and 0s. These strings fail the unpredictability test (see above). If probability requires randomness, probability is not identical to frequency in the long run. It seems unreasonable, e.g. to assign .5 to probability of a 1 at a given point in the sequence because the frequency of 1s in the long run is .5, if the chance of getting a 1 at any point in the sequence can be determined exactly to be 1 or 0.

29. Shannon entropy can be reduced by a passive filter; this is impossible for the entropy used in physics [11, p.161]. Shannon's use of "entropy" has caused more confusion for students in the field than any event since Clausius' victory over Tait. See, however, [78, p.18], for a different view of the relation between Shannon entropy and physical entropy.

30. The confusions and resulting disputes are over definitions of terms, but to view them as *merely* definitional obscures the substantive issues involved, namely, the relation of information theory to physical theory, and all that it implies about the nature of computation and physical dynamics.

31. A different, but still conservative approach, is to compare the budgets qualitatively, but to avoid directly unifying the concepts or extending the concepts in any way [138, 123]. Tom Schneider's [124] quantitative unification of the information and entropy budgets is conservative in its rejection of conceptual unification and any consequent possibility of conceptual extension.

32. This definition is more specific than the standard definition of work in equilibrium thermodynamics, in which work is a force applied through a distance. Requiring a vector is reasonable, since undirected force can not really do work. However, this changes the units of work, since energy is not a vector. Interestingly, Schrödinger [125] says that he would have used *exergy*, the measure of energy available in a system to do work, to ground NPI, except that energy already had specific associations that might be confusing, so he used negentropy instead. This is remarkable, since exergy and entropy do not have the same dimensions.

33. Our statement here involves some stylistically sensitive matters. Brillouin [11, p.152] refers to physical information as bound information but, in the light of the later distinction between intropy and enformation (see below), we will avoid this term (since in one obvious sense intropy, being unconstrained by the system, is not bound). Brillouin defines bound information as a special case of free information, which is abstract, and takes no regard of the physical significance of possible cases. By contrast, we would say that bound information occurs when the possible cases can be regarded as the complexions of a single physical system. We may consider algorithmic bound information as an extension of this idea requiring that the strings of binary digits representing the information are isomorphic map-

pings of physical structures, where alternate physically permissible structures are the relevant complexions. This definition also has the consequence that the physical interpretation of the two formulations of physical information are not only mathematically, but also physically, equivalent. The basis of the mathematical equivalence was proposed by Kolmogorov, but given by Ingarden, [78]. Although physical entropy and physical information are complementary, and hence opposite in sign, they measure the same fundamental physical notion, hence our apparent difference with Ingarden et al is a matter of presentation rather than a substantive difference. This stylistic difference has caused some confusion in the past about the interpretation of NPI, especially as used by Brillouin. It is important to recognise that there is no underlying difference of fact.

34. With respect to the need to interpret the principle in relation to the system and environment under consideration, the situation is exactly parallelled by that for energy and momentum. Energy forms are specific to each system — electrical circuits store and conduct electrical energy but radiation and thermal energy are typically ignored, unless the system includes light bulbs, electrolysis and the like, while the potential energy in spring extension is of a quite different character to that in chemical density and capacitance voltage, and so on. Potential energy is determined relative to a specific environment, e.g the potential energy of an object at temperature T depends on the temperature of its surrounds, the kinetic energy of a particle is a function of its velocity which must be specified relative to a frame of reference. In our case, by referring information to the system environment we avoid the need to define some absolute reference point where all constraints of any kind are relaxed, which is not obviously a well defined condition. (While conservative systems might support some such notion as energy of separation of system components to infinity, this is not obviously well defined when non-linear intra- and inter-interactions are involved.) And just as there are very different formulae for all these forms of potential energy in different systems, so too are there for forms of entropy and information. But whereas all these formulations can be uniquely connected through the principles of energy conservation (and similarly for momentum), there is currently no equivalent uniqueness achievable with formalisms for entropy and information. Perhaps a canonical version of NPI that can be applied more or less mechanically will emerge some day, and Ingarden [78] represents progress here, but with the proliferation of disparate "entropies" and versions of information currently afoot, this is unlikely to happen soon.

35. There are strong reasons to believe that this is logically impossible in a physicalist world [37].

36. This is not quite as simple as Szillard's case (see [11, pp.176ff]), which uses only one molecule!

37. NPI is assumed throughout, as is the impossibility of a Maxwellian demon [11, 4, 37]. Szillard's original argument makes the connection to work more obvious by using a molecule pushing on a cylinder in a piston, but more general arguments by Bennett and Collier examine (in different ways) the computational problem the demon is supposed to solve. The connection to work is implied by thermodynamics and NPI. Szillard used thermodynamics explicitly, but NPI only implicitly, which meant that his exorcism of the demon could not be general. Denbigh [48] argue that information is not required for the exorcism, since thermodynamics can be used in each instance. It seems to have escaped them that proving this requires something at least as strong as NPI. The problem of Maxwell's demon is especially important because it forces us to be explicit about the relations between mental and physical activity. A demon that could store information in some non-physical form could perform its sorting job [34].

38. See [36]. For endothermically produced systems S (e.g. smelted steel objects) when enformation decays additional energy will be obtained from S as the cohesive constraints on S are relaxed or disrupted. This process can be accelerated through the addition of triggering energy (e.g. ignition in chemical or nuclear processes), but will occur eventually if we wait long enough, and the temperature is above absolute zero. The released energy will be $\int T\Delta\varepsilon$, and will eventually degrade into heat, indicating that decaying enformation is a form of intropy. There may be cases in which enformation remains even at equilibrium, such as enformation produced by an exothermic process; in such cases the enformation cannot decay, and is energetically inaccessible. A possible example is the enformation in all the protons of the universe, assuming protons do not decay spontaneously, and that there is an insufficient supply of antiprotons or potential antiprotons to annihilate all of them. This would seem to include there remaining enformation even at absolute zero, raising interesting, unresolved issues about the status of the so-called Zeroth Law of thermodynamics, which we do not pursue here.

39. Since one obvious information basis to consider is a complete microscopic description of a system, we note that behind this statement lies the vexed issue of a principled resolution of the relations between statistical mechanics and thermodynamics that respects the irreversibility of the latter despite the reversibility of the former. While we think that the analyses offered here represent a small step toward greater clarity about this complex issue, we do not pursue this issue here.

40. Archimedes lever with which he could move the world, like any other machine, must have a specific form: it must be rigid, it must be long enough, there must be a fixed fulcrum, and there must be a force applied in the right direction. If any of these are lacking, the lever would not work. No amount of energy applied without regard to the form in which it is applied can do work, except by accident. Collier [36] who introduced the intropy/enformation distinction, did not provide

a differential form for enformation, unlike intropy and entropy. There seems no principled reason why a differential form could not be introduced but it would be of dubious relevance since, unlike the latter, it is not intrinsically a thermodynamic property of an ensemble. It has been noted [36] that the statistical measure is most suited to situations where the information is specified over an ensemble, while the complexity measure is best suited to single cases. We can access the intropy only by measuring changes of state (since we cannot directly measure the components of the ensembles, most being imaginary), but we can determine the enformation directly from the state.

41. There is nothing arbitrary about these system-relative distinctions, each is grounded in the system dynamics. Relational properties, like intropy, entropy and enformation, necessarily produce relativised applications across relationally distinct contexts, e.g. S and S' here, and it is an error (albeit a common one) to equate this to relativism, which is the absence of any principled basis for distinguishing conflicting claims across contexts. Also, in the presence of system-relative distinctions, some mistake our free choice of which system to model, e.g. S or S', for an idealist or subjectivist component to system dynamics itself, neglecting the objective dynamical grounds for relative attributions. Salthe [118, 119] and Kampis [79], for example, both seem to make this mistake. Possibly it can be traced back to Ashby. Finally, one can change a relational property of a system by re-relating system components, i.e. without bringing anything new into the system, but one cannot do this with non-relational properties, a change in those alters at least the quantity of their kind of "stuff". For this reason it can be misleading to speak of creating or transforming relational properties, but where system-relative distinctions are grounded in the system dynamics the usage is a principled one and where context is clear no ambiguity should result. Thus we say, for example, that connecting S to P creates the S' intropy expressed by G' because there is a principled dynamical difference between S, P mutually isolated and S, P physically connected in a way that allows heat transfer (and the particular way connected also matters for the resulting process dynamics).

42. All enformation except perhaps the enformation in some fundamental particles, like protons, will eventually decay, which means that at some temporal scale all, or at least most, enformation behaves as intropy (note 38). The scale is set by natural properties of the system in question. Specifically, the extent of the cohesion of the system implies a natural scale.

43. A complete physical specification would amount to a maximally efficient physical procedure for preparing the system, S, in the macrostate in question from raw resources, R [36]. Furthermore, the procedure should be self-delimiting (it halts when S is assembled, and only when S is assembled). The information content of this specification is just I_P plus any intropy that must be dissipated in the process. The latter is called the *thermodynamic depth* of the state of the system, and

is equal to $H_{\mathrm{ACT}}(R) - H_{\mathrm{ACT}}(S)$ if there are no practical restrictions on possible physical processes. The algorithmic complexity analogue of thermodynamic depth is the complexity decrease between the initial and final states of a computation (through memory erasure). This quantity is often ignored in algorithmic complexity theory, but see [5, 37, 54], where the authors would hold that the analogy is a physical identity.

44. There is one further terminological issue concerning physical information that should be noted. By NPI, the disordered part of the system does not contain information (because it cannot contribute to work), but the information required to specify the complete microstate of the system is equal to the information in the macrostate plus the information required to specify the disordered part. Layzer [90, 91] speaks of the information required to specify the disordered part as the "microinformation" of microstates, as if the information were actually *in* the microstate. This information can do work only if it is somehow *expressed* macroscopically. For this reason, we prefer to regard unexpressed microinformation as a form of *potential information* [56, 33, 13]. Expressed information in the form of enformation is sometimes called *stored information* [56, 13]. Potential information can also be *directly expressed* as intropy, e.g. in the Brownian motion of a particle, as opposed to at the expense of enformation, e.g. when micro fluctuations disrupt structure. Although expression as intropy is physically possible, it cannot be physically controlled [37]. Control of this process would imply the existence of a Maxwellian demon.

45. Ulanowicz [133] refers to this quantity as system overhead Φ, but an overhead is a cost of some kind and we do not follow this terminology. There is already defined a very different quantity, and genuine cost, computational overhead, which is a technical nuisance to be worked around. The intropy is a measure of the variability of the unconstrained part of the system, and is better refered to simply as the *variability*.

46. An autonomous system may use its self-organisation potential to compensate for assaults on its integrity, though this would typically be achieved through reorganisation, but must use it to achieve genuinely new higher order organisation than it currently possesses — see Section 7.

47. Note that if NPI is not assumed, further argument is required to establish that the large C_I of a gas cannot be used to control its behaviour. The argument would have to be on a case by case basis, similar to the situation if we accepted the Denbighs' [48] rejection of the relevance of information theory to Maxwell's demon. It might seem obvious that type 2 systems cannot support autonomy or anticipation, but the reasons why are non-trivial.

48. Assuming a protein can change from a denatured, linear form to an active, folded form reversibly, the folded form contains more conformational information

than the linear form. For I_P to remain constant, the additional conformational information would have to be implicit in the dynamical information of the linear form, and vice versa. In practice, protein folding is not reversible.

49. For a metaphysical argument for the same result, not restricted to conservative systems, see [38].

50. Despite this, all of the early applications of modern physics were to linear or approximately linear systems, essentially type 2 or type 1 systems. The fact that these early applications have served as Kuhnian exemplars for later applications [85] has led to an overly simplified view of dynamics. This is ironic, since Kuhn's own theory of science uses a highly simplified dynamics that misleads as it informs ([64, 65], see note 23).

51. Ironically, even relatively minor dissipation, like tidal dissipation, can lead to phase spaces with multiple point attractors (see below) in which the boundaries between attractors are chaotic. In systems in which the trajectory enters chaotic regions in phase space (the Sun-Mercury system is likely one of these) the method of linear approximations can give only probabilities of capture by one attractor or another.

52. It does not. DNA almost certainly exhibits high order subtle redundancies (see below) imposed by developmental, environmental and self-organised constraints through the process of evolution [33, 13].

53. For a review of nonlinear systems and chaos, see [132].

54. See Section 6 below. Penrose sometimes talks in *The Emperor's New Mind* as if the tilings were not effectively computable. If not, then the argument in the previous paragraph concerning the capacities of aperiodic crystals implies that they cannot control in any interesting way. If the tilings are non-computable, then they are immediately irrelevant to the control aspects of mind unless somehow we can non-computably identify suitable partial functions that are computable, and are relevant to control problems. We do not see how this is supposed to work. More likely, Penrose thinks that *finding* a tiling, given a set of tiles, is non-computable, but we can somehow do it, and that finding aperiodic "patterns" like tilings is useful for control problems. The last is at least plausible, since the possible tilings are a very specific subset of a myriad of possible combinations, yet they integrate local and global constraints. As we have previously noted, this sort of integration is a characteristic of organic systems. If, on the other hand, evolution has selected effectively computable physiological and behavioural tiling procedure analogues from the myriad possibilities mutation and self-organisation put forth so as to "solve" specific global/local integration problems, we need not assume any mental capacities that are not representable as computations.

55. Cf. note 47 and text. A common case is condensation after rapid expansion of the volume containing a gas. The treatment of gravitational condensation may be more complicated, see [91].

56. In practice, dissipation will set in much faster than organisation forms, especially since there is nothing like minimisation of entropy production to guide the formation of physical information. Even in a non-expanding system, ordered structures *could* appear spontaneously. However, there would be nothing to drive their formation, and chance interactions are more likely to disorganise the system rather than organise it (cf. also [53, appendices]. Any new organisation appearing in a conservative system, then, could only be attributed to chance or supernatural causes [37]. For all practical purposes, self-organisation requires dissipation. See also notes 14 and 17.

57.This is a strictly mathematical decomposition. Physical decomposability is not required. It is important not to confuse orders with levels (see below).

58. So complex organisation neither requires, nor is required by, chaos and the organisation of living systems is typically not chaotic. While the role of chaos in living processes is an empirical and open issue at the present time, that chaos derives from the extreme sensitivity of its dynamical trajectories to initial conditions provides a good reason why it will probably not be common in adaptable creatures where resources are limited and adaptation typically must be fast and accurate (though it may be used selectively to achieve sensitive control, [127, 38]).

59. Some adjustments are required to the definition to get a reasonable value of depth for finite strings. We want to rule out cases in which the most compressed program to produce a string is slow, but a slightly longer program can produce the string much more quickly. To accommodate this problem, the depth is defined relative to a significance level s, so that the depth of a string at significance level s is the time required to compute the string by a program no more than s bits longer than the minimal program.

60. Since computation is a formal concept, while time is a dynamical concept, it is not completely clear how we can get a dynamical measure of computation time. Jonathan Smith [128] has shown that the formal analogue of temperature in informational systems has the dimensions of inverse time (i.e. a rate). Together with NPI, this *may* give us a quantitative measure for the physical rates involved in the informational dynamics of systems. Generally, the minimal assembly time of a system will be less than the expected assembly time for assembly through random collisions, which we can compute from physical and chemical principles. Maximally complex systems are an exception, since they can be produced only by comparing randomly produced structures with a non-compressible template. Because of their compressibility, organised systems can be produced by finding a compressed form, and producing the structure from that. This is more efficient

than checking randomly produced structures, and has clear implications for the relevance of evolutionary processes as effective search-and-construction processes. Cf. note 43.

61. The relative logical depth construction captures something important to these systems. Too much organisation is detrimental to living systems because it is too constraining and/or too costly to maintain (cf. below) and, for the same reasons, so also is too much redundancy. It is also worth noting that it appears to take less time to assemble something from modules than to create functionally equivalent organisation unimodularly. We should then expect intermediate organisation in AAAR systems, irrespective of type. Nonetheless, while the relative logical depth construction is a major move in the direction of specifying organisation, it is unclear whether it will be sufficient. In any case, it needs a clear dynamical interpretation to be applied to hierarchial and control organisation.

62. For example, computer programs using a serial von Neumann architecture can be implemented conservatively, at least in principle [54, 3]. Such a computer could control another conservative system, or could control states of its own subsystems. DNA is neither formed nor maintained under conservative conditions, but, supposing it could be, DNA could control the growth and development of an organism in much the way that the contemporarily popular computer program analogy of DNA suggests. The linear and modular structure of the DNA molecule lends support to the program metaphor, and much of our present practical understanding of the dynamics of DNA transcription and protein interactions depends only on conformational information, further supporting the computational model. Whether or not this model is ultimately workable, conservative systems with enough logical depth are capable of anticipation as Bickhard has characterised it [8, 9].

63. Depth may increase spontaneously relative to initial conditions in many conservative systems. In an ideal gas at equilibrium, for example, correlations among molecules resulting from collisions increases the depth of the microstate relative to the initial conditions, but there is no increase in the absolute depth, since the initial condition is random [5]. In a non-equilibrium gas, the absolute depth can increase by a similar process, if the increase is not obliterated by noise, since the initial condition is not random. The no-noise assumption is not usually realistic, since it implies no dissipation under non-equilibrium conditions. Therefore we are unlikely to find conservative systems that increase their absolute depth.

64. More widely affirmed than argued, the argument to this general conclusion runs roughly as follows (see especially [122, 123]: The local time rate of entropy production = dissipation rate = local thermodynamic potential, so the spatial gradient of the potential = the spatial gradient of the time rate of entropy production, so total time rate of entropy production = spatial integration of thermodynamic potential = total potential difference across the system; but the spatial gradient

of the potential = generalised force and, by generalised laws of motion, systems re-arrange until force is minimised, so systems rearrange until the spatial potential gradient is minimised, so systems rearrange until the total potential difference is minimised; but all possible potential distributions for a given system have the same minimum entropy production rate, so systems rearrange until the total entropy production is minimised. An equivalent principle is that exergy loss is minimised = minimisation of input work to maintain the system in a steady state. In the case of Bénard cells, for example, less work is needed to maintain convection than conduction. This argument has a more precise version for near-to-equilibrium systems, but is claimed to extend to all systems. Though this reasoning is plausible, we do not find it transparent (especially at the last step above) and look forward to future improvements in foundations and rigour. Nonetheless, the principles seem well founded empirically thus far, and are widely accepted.

65. For a detailed treatment of the dynamics underlying the transition, see [41]. A number of other examples have been studied with varying degrees of detail, but the general pattern has been well verified, both in theory and fact.

66. Near-to-equilibrium systems are ones in which the local statistical fluctuations are larger than or equal to the local gradient of intensive state variables. They can be treated pretty much like equilibrium systems, but they exhibit some novel properties. These properties result from the entropy production from dissipation, characterised by the specific (local) rate of entropy production. In particular, they can self-organise.

67. Compare here to the fixed stoney bed/bank river system a river running through mud; the river reconstructs its muddy boundaries, its bed and bank, as a result of its own dynamics, and so alters its own flow patterns, which in turn alter its subsequent impact on its bed and banks, and so on.

68. This is connected to our reservation about the adequacy of relative logical depth to capture a full concept of dynamical organisation, see Section 6 above. There are interesting further questions as to whether there are additional or different resources intrinsically required for cognition and to what extent all cognitive systems, living and artificial, have the same requirements. It is our view that at least all biologically based cognitive systems are essentially similar in these system respects, and that the basic requirements are met by AAA systems — see Conclusion.

69. This is a complex relationship, much of which is implicit for the adaptive systems concerned. Roughly, the meaning of a signal for such a system is what they do with it, its anticipative content. If for system S signal I initiates action a then the meaning of I for S is "This is an a-appropriate environment". Note that on this account the information in a signal is a function of its downstream modulation of response, as opposed to the traditional upstream sender state, a shift

which solves many problems while arising naturally from systems operation. The achievement of a detached characterisation of sender states is a sophistication that only arises with the co-ordinated extension of higher order anticipative control. This notion satisfies the Shannon/Weaver definition of signal information since signals appropriately reduce uncertainty of system response and to that extent can be used to control response. For a more detailed treatment of information in a signal as a function of its downstream modulation of response (as opposed to the traditional upstream sender state) see [8, 26, 30]. For a treatment of information as a sign, following suggestions by Carl von Weisacker, see [86].

70. In this way we provide a principled basis in our analysis for the Ghiselin/Hull notion of species as individuals, [57, 58, 75, 76, 77]. However, in one crucial respect they cannot have the full autonomy that individual phenotypes enjoy, namely to receive feedback from their (temporally extended and variegated) environment which can be evaluated for its effectiveness in supporting autonomy. By contrast with reproductive capacity, there is at best a very partial equivalent to individual regenerative capacity for lineages because the internally organised parts of themselves are temporally separated. Environmental feedback can only be evaluated by individual phenotypes and only derivatively by lineages, it cannot be evaluated by the lineage *acting as an integrated individual*.

71. Such self-organising forces may include speciation under segregation of gene flow in a population resulting from geographical isolation and/or social differentiation, other forms of population isolation permitting group selection, fixation of group level properties, etc. [103, 104] on isolating mechanisms; [13, p.208,pp.215–218], [14, 119, 61]. Of course, there may also be selection for speciation, as when hybrids are selectively removed in favour of the pure forms (homozygotes), whether the selection is internal (e.g. hybrids refuse to mate, or cannot produce offspring), social (e.g. hybrids killed by pures) or environmental (e.g. hybrid predation higher). At present, there has been little empirical work to quantify the strength of selection in relation to self-organisation; it is merely presumed to be dominant. It is intriguing, though, that members of several ancient, long-lived, widely and homogenously distributed pine species do not seem to be optimally adapted to local conditions [100, 101, 102].

72. A variant on vicariant selection is the neural selection of Edelman [50]. If, as with concepts and principles, we consider the gene type, rather than the individual gene instances or tokens, then we may regard phenotypes as vicariant life forms on which selection acts as the gene types search for better vicariants; then vicariance is not unique to psychological adaptation, but the use of neural resources for constructing vicariants, e.g. logical constructions, is.

73. But putting together a coherent, well founded account of intentionality focused on system organisational design is a complex task; its beginnings can be found in

marrying system-constructed signal semiotics or significance (see note 69) with system organisation as discussed here and in [21, 22, 24] and applying to it an analysis of system heuristics [26], [27]–[30]).

Bibliography

1. P. W. Atkins, *The Second Law: Energy, Chaos, and Form*, New York, Scientific American Library, 1994.
2. M. Baatz, *Pleiotropy and the evolution of adaptability: a population genetic contribution to a systematic understanding of evolution*, in Collier and Siegel-Causey, 1999.
3. C. H. Bennett, IBM Journal of Research and Development **17**, 525 (1973). Reprinted in Lef and Rex 1990.
4. C. H. Bennett, International Review of Theoretical Physics **21**, 905 (1982). Reprinted in Lef and Rex 1990.
5. C. H. Bennett, C. H. 1985. "Dissipation, information, computational complexity and the definition of organization", in: D. Pines, ed., *Emerging Syntheses In Science*, Proceedings of the Founding Workshops of the Santa Fe Institute, Redwood City, Calif., Addison West Publishing Company, 1985.
6. C. H. Bennett, IBM Journal of Research and Development **32**, 16 (1988). Reprinted in Lef and Rex 1990.
7. A. M. Bertgache and K. Ruiz-Mirazo, Biosystems **49**, 45 (1999).
8. M. H. Bickhard, Journal of Experimental and Theoretical Artifical Intelligence **5**, 285 (1993).
9. M. H. Bickhard and L. Terveen, *Foundational issues in artificial intelligence and cognitive science: impasse and solution*, New York, Elsevier, 1995.
10. V. Braitenberg, *Vehicles*, Cambridge, Mass., MIT Press, 1984.
11. L. Brillouin, *Science and Information Theory*, second edition, New York, Academic Press, 1962.
12. D. R. Brooks, J. D. Collier, B. A. Maurer, J. D. H. Smith, and E. O. Wiley, Biology and Philosophy **4**, 407 (1989).
13. D. R. Brooks, and E. O. Wiley, *Evolution as Entropy: Toward a Unified Theory of Biology*, second edition, University of Chicago Press, 1988.
14. D. R. Brooks and D. A. Maclennan, *Phylogeny, ecology and behaviour*, University of Chicago Press, 1991.
15. R. A. Brooks, Artificial Intelligence **47**, 139 (1991).
16. R. M. Burian, "Adaptation: Historical perspectives", in: Keller and Lloyd, 1992.
17. S. R. Caplan and A. Essig, *Bioenergetics and linear nonequilibrium thermodynamics*, Cambridge, Harvard University Press, 1983.
18. D. T. Campbell, "Downward causation in hierarchically organized biological systems", in: F. J. Ayala and T. Dobzhansky, eds., *Studies in the Philosophy of Biology*, New York, Macmillan, 1974.
19. D. T. Campbell, "Evolutionary epistemology", in: P. A. Schilpp, ed., *The philosophy of Karl Popper*, La Salle, Open Court, 1974.
20. C. Cherniak, *Minimal rationality*, Cambridge, Bradford/MIT Press, 1986.
21. W. D. Christensen, Biology and Philosophy **11**, 301 (1996).
22. W. D. Christensen, *An interactivist-constructivist approach to adaptive intelligence and agency*, Doctoral Thesis, University of Newcastle, Australia, 1999.
23. W. D. Christensen, J. D. Collier, and C. A. Hooker, "Autonomy, adaptiveness and anticipation: towards foundations for life and intelligence in complex, adaptive, self-organising systems", circulated draft.
24. W. D. Christensen, J. D. Collier, C. A. Hooker, "Adaptiveness and adaptation: A new autonomy-theoretic analysis and critique", submitted.
25. W. D. Christensen, C. A. Hooker, "Very simple minds: Towards a general theory of intelligent systems", in preparation.

26. W. D. Christensen and C. A. Hooker, "From Cell to Scientist: Toward an organisational theory of life and mind", in: J. Bigelow, ed., *Our Cultural Heritage*, Australian Academy of Humanities, University House, Canberra, 1998.
27. W. D. Christensen and C. A. Hooker, "Organised interactive construction: The nature of autonomy and the emergence of intelligence", in: A. M. Bertgache and A. Etxebberria, eds., *The Contribution of Artificial Life and the Sciences of Complexity to the Understanding of Autonomous Systems, Communication and Cognition*, Special Edition, 1999.
28. W. D. Christensen and C. A. Hooker, "Autonomy, Heuristic Policies and Self-Directed Learning", submitted.
29. W. D. Christensen and C. A. Hooker, "The organisation of knowledge: beyond Campbell's evolutionary epistemology", *Proceedings, Sixteenth Biennial Meeting of the Philosophy of Science Association 1998, Philosophy of Science*, Special Edition, 1999.
30. W. D. Christensen and C. A. Hooker, "The organisation of knowing: Autonomy, evolution, and evolutionary epistemology", in: [66].
31. P. M. Churchland, *A Neurocomputational Perspective*, Cambridge, Bradford/MIT Press, 1989.
32. P. M. Churchland, *The Engine of Reason, the Seat of the Soul*, Cambridge, MIT Press, 1995.
33. J. D. Collier, Biology and Philosophy **1**, 5 (1986).
34. J. D. Collier, Biology and Philosophy **3**, 501 (1988).
35. J. D. Collier, "Supervenience and reduction in biological hierarchies", in: M. Matthen and B. Linsky, eds., Philosophy and Biology: Canadian Journal of Philosophy, Supplementary Volume **14**, 209 (1988).
36. J. D. Collier, "Intrinsic information", in: P. Hanson, ed., *Information, Language and Cognition, Vancouver Studies in Cognitive Science, Vol. 1*, Oxford, University of Oxford Press, 1990.
37. J. D. Collier, Studies in the History and Philosophy of Science **21**, 257 (1990).
38. J. D. Collier, Symmetry: Culture and Science **7**, 247 (1996).
39. J. D. Collier, "Information increase in biological systems: how does adaptation fit?", in: G. van der Vijver, S. N. Salthe and M. Delpos, eds., *Evolutionary Systems*, Dordrecht, Reidel, 1998.
40. J. D. Collier, "Incorporating adaptation into evolution as entropy", in [42].
41. J. D. Collier, S. Banerjee and L. Dyck, "A non-equilibrium perspective linking development and evolution", in: [42].
42. J. D. Collier and D. Siegel-Causey, eds., *Between Order and Chaos: Studies in Non-Equilibrium Biology*, submitted.
43. M. Conrad, Systems Research **10**, 3 (1993).
44. M. Conrad, "Towards high evolvability dynamics", in: G. van der Vijver, S. N. Salthe and M. Delpos, eds., *Evolutionary Systems*, Dordrecht, Reidel, 1998.
45. F. Crick, *The Astonishing hypothesis: The scientific search for the soul*, New York, Scribners, 1994.
46. R. Cummins, *The nature of psychological explanation*, Boston, Bradford/MIT, 1983.
47. R. Dawkins, *The Blind Watchmaker*, New York, W. W. Norton, 1987.
48. K. G. Denbigh and J. S. Denbigh, *Entropy in Relation to Incomplete Knowledge*, Cambridge University Press, 1985.
49. C. Dyke, *The Evolutionary Dynamics of Complex Systems: A Study in Biosocial Complexity*, New York, Oxford University Press, 1988.
50. G. M. Edelman, *Neural Darwinism: The Theory of Neuronal Group Selection*, New York, Basic Books, 1987.
51. M. Eigen and P. Schuster, *The Hypercycle: A Principle of Natural Self- Organization*, Berlin, Springer, 1979.
52. N. Eldredge, *Unfinished Synthesis: Biological Hierarchies and Modern Evolutionary Thought*, New York, Oxford University Press, 1985.
53. P. Fong, *The Unification of Science and Humanity*, Stillwater, New Forums Press, 1996.
54. E. Fredkin and T. Toffoli, International Journal of Theoretical Physics **21**, 219 (1982).

55. A. Garfinkel, "The slime mould Dictyostelium as a model of self-organisation in social systems", in: E. F. Yates, ed., *Self-organizing Systems: the Emergence of Order*, New York, Plenum Press, 1987.
56. L. L. Gatlin, *Information Theory and the Living System*, Columbia University Press, New York, 1972.
57. M. Ghiselin, Systematic Zoology **23**, 536 (1974).
58. M. Ghiselin, Biology and Philosophy **2**, 127 (1987).
59. R. Gilmore, *Catastrophe theory for scientists and engineers*, New York, Wiley, 1981.
60. H. Goldstein, *Classical Mechanics*, 2nd edition, Reading, Addison-Wesley, 1980.
61. B. Goodwin, *How the Leopard Changed its Spots: The Evolution of Complexity*, London, Weidenfeld and Nicolson, 1994.
62. K. Hahlweg and C. A. Hooker, "Evolutionary epistemology and philosophy of science", in: [63].
63. K. Hahlweg and C. A. Hooker, *Issues in Evolutionary Epistemology*, Albany, State University of New York Press, 1989.
64. W. E. Herfel and C. A. Hooker, "Cognitive dynamics and the development of science" in: D. Ginev and R. S. Cohen, eds., *Issues and Images in the Philosophy of Science*, Dordrecht, Kluwer, 1997.
65. W. E. Herfel and C. A. Hooker, "From formal machine to social colony: toward a complex dynamical philosophy of science", forthcoming in: M. L. Dalla Chiara, D. Mundici, J. van Bentham, eds., *Proceedings of the 10th International Congress of Logic, Methodology and Philosophy of Science*, Boston, Kluwer, 1999.
66. C. Heyes and D. Hull, eds., *Festschrift for Donald Campbell*, submitted.
67. J. H. Holland, *Adaptation in Natural and Artificial Systems*, Cambridge, MIT/Bradford, 1992.
68. C. A. Hooker, Dialogue XX, 496 (1981).
69. C. A. Hooker, Biology and Philosophy **9**, 197 (1992).
70. C. A. Hooker, *Reason, Regulation and Realism: Toward a Naturalistic, Regulatory Systems Theory of Reason*, Albany, State University of New York Press, 1995.
71. C. A. Hooker, Philosophy and Psychology **8**, 287 (1995).
72. C. A. Hooker, "Toward a naturalised cognitive science", in: R. Kitchener and W. O'Donohue, eds., *Psychology and Philosophy*, London, Sage, 1996.
73. C. A. Hooker, H. B. Penfold, and R. J. Evans, British Journal for the Philosophy of Science **43**, 517 (1992).
74. J. Horgan, Scientific American **272**, no. 6, 77 (1995).
75. D. L. Hull, Systematic Zoology **25**, 174 (1976).
76. D. L. Hull, Philosophy of Science **45**, 335 (1978).
77. D. L. Hull, Annual Review of Ecology and Systematics **11**, 311 (1980).
78. R. S. Ingarden, A. Kossakowski, and M. Ohya, *Information Dynamics and Open Systems: Classical and Quantum Approach*, Dordrecht, Kluwer, 1997.
79. G. Kampis, *Self-Modifying Systems in Biology and Cognitive Science: A New Framework for Dynamics, Information and Complexity*, Oxford, Pergamon Press, 1991.
80. S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford, Oxford University Press, 1993.
81. E. F. Keller and E. A. Lloyd, eds., *Keywords in Evolutionary Biology*, Cambridge, Harvard University Press, 1992.
82. J. Kestin, *A Course in Thermodynamics*, Waltham, Blaisdell, 1968.
83. A. N. Kolmogorov, Problems of Inform. Transmission **1**, 1 (1965).
84. A. N. Kolmogorov, IEEE Transactions on Information Theory 14, 662 (1968).
85. T. Kuhn, *The Structure of Scientific Revolutions*, 2nd edition, University of Chicago Press, 1970.
86. B-O. Küppers, *Information and the origin of life*, Cambridge, MIT Press, 1990.
87. R. Landauer, IBM J. Res. Dev. **5**, 183 (1961); reprinted in [92].
88. R. Landauer, Phys. Scr. **35**, 88 (1961), reprinted in [92].
89. P. T. Landsberg, Physics Letters A **102**, 171 (1984).

90. D. Layzer, Scientific American **233**, 56 (1975).
91. D. Layzer, *Cosmogenesis: The Growth of Order in the Universe*, New York, Oxford University Press, 1990.
92. H. S. Lef and A. F. Rex, *Maxwell's Demon: Entropy, Information, Computing*, Princeton University Press, 1990.
93. M. Li and P. Vitànyi, "Kolmogorov complexity and its applications", in: J. van Leeuwen, ed., *Handbook of Theoretical Computer Science*, Dordrecht, Elsevier, 1990.
94. E. Mach, *The Science of Mechanics*, Lasalle, Open Court, 1960.
95. C. B. Martin, Pacific Philosophical Quarterly **65**, 239 (1983).
96. K. Matsuno, *Protobiology: Physical Basis of Biology*, Boca Raton, CRC Press, 1989.
97. K. Matsuno, ed., Contemporary Philosophy **24** (11) (1996).
98. K. Matsuno and S. N. Salthe, Biology and Philosophy **10**, 309 (1995).
99. H. Maturana and F. Varela, *Autopoiesis and cognition*, Dordrecht, Reidel, 1980.
100. C. Mátyás, "Genetic and ecological constraints of adaptation", in: *Forest Genetics: Breeding and Physiology of Woody Plants*, Moscow, The State Forest Committee of the USSR, 1989.
101. C. Mátyás, "Climatic adaptation of trees: rediscovering provenance tests", Euphytica, in press.
102. C. Mátyás and C. Yeatman, Silvae Genetica **41** 370 (1992).
103. E. Mayr, *The Growth of Biological Thought: Diversity, Evolution and Inheritance*, Cambridge, Harvard University Press, 1982.
104. E. Mayr, *Towards a New Philosophy of Biology*, Cambridge, Harvard University Press, 1988.
105. G. Nicolis and I. Prigogine, *Self-Organization in Non-Equilibrium Systems: From Dissipative Structures to Order Through Fluctuations*, New York, John Wiley and Sons, 1977.
106. M. H. Nitecki, ed., *Evolutionary Progress*, University of Chicago Press, 1988.
107. H. H. Pattee, "Physical theories of biological co-ordination", in: M. Grene and E. Mendelsohn, eds., *Topics in the Philosophy of Biology*, Dordrecht, Reidel, 1976.
108. R. Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, New York, Oxford University Press, 1989.
109. H. C. Plotkin, Biology and Philosophy **2**, 87 (1987).
110. H. C. Plotkin, Biology and Philosophy **6**, 481 (1991).
111. H. Plotkin, *Darwin Machines*, Cambridge, Harvard University Press, 1994.
112. K. Pribram and M. Gill, *Freud's 'Project' Re-assessed*, London, Hutchinson, 1976.
113. I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes*, 2nd edition, New York, John Wiley and Sons, 1961.
114. I. Prigogine, *From Being to Becoming*, San Francisco, Freeman, 1980.
115. I. Prigogine and I. Stengers, *Order out of Chaos*, New York, Bantam Books, 1984.
116. Z. Pylyshyn, *Computation and Cognition: Toward A Foundation for Cognitive Science*, Cambridge University, 1984.
117. R. Rosen, "Organisms as Causal Systems Which Are Not Mechanisms: An Essay into the Nature of Complexity", in: R. Rosen, ed., *Theoretical Biology and Complexity*, London, Academic Press, 1985.
118. S. N. Salthe, *Evolving Hierarchical Systems*, New York, Columbia University Press, 1985.
119. S. N. Salthe, *Development and Evolution: Complexity and Change in Biology*, Cambridge, MIT Press, 1993.
120. S. N. Salthe, Ludis Vitalis **3**, 15 (1995).
121. J. C. Schanck and W. C. Wimsatt, "Generative entrenchment and evolution", in: A. Fine and P. K. Machamer, eds., *PSA 86, Volume 2*, East Lansing: Philosophy of Science Association pp. 33–60, 1987.
122. E. Schneider and J. J. Kay, "Order from disorder: the thermodynamics of complexity in biology", in: M. P. Murphy and L. A. O'Neil, eds., *What is Life: the Next Fifty Years*, Cambridge University Press, 1994.
123. E. Schneider and J. J. Kay, "Life as a manifestation of the Second Law of Thermodynamics", in [42].

124. T. S. Schneider, "An equation for the Second Law of Thermodynamics", Word Wide Web URL: ¡http://www-lmmb.ncifcrf.gov/ toms/paper/secondlaw/index.html¿
125. I. Schrödinger, "What is Life?", reprinted in *What is Life? And Mind and Matter*, Cambridge University Press, 1944.
126. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, Urbana, University of Illinois Press, 1949.
127. C. A. Skarda and W. J. Freeman, Behavioural and Brain Sciences **10**, 161 (1987).
128. J. D. H. Smith, "Mathematical approaches to the unified theory of biology", in [42].
129. T. Smithers, "Are Autonomous Agents Information Processing Systems?", in: L. Steels and R. A. Brooks, eds., *The Artificial Life Route to 'Artificial Intelligence': Building Situated Embodied Agents*, Hillsdale, Lawrence Erlbaum Associates, 1995.
130. R. Sperry, *Science and Moral Priority: Merging Mind, Brain and Human Values*, New York, Columbia, 1983.
131. K. Sterelny, K. C. Smith and M. Dickison, Biology and Philosophy **11**, 377 (1996).
132. H. B. Stewart, and J. M. Thompson, *Nonlinear Dynamics and Chaos*, Chichester, Wiley, 1986.
133. R. E. Ulanowicz, *Growth and Development: Ecosystems Phenomenology*, New York, Springer Verlag, 1986.
134. R. E. Ulanowicz, "Boundaries on the complexity of evolving networks; a window on vitality", in [42].
135. C. H. Waddington, *Principles of Development and Differentiation*, New York, MacMillan, 1966.
136. P. J. Ward, "A model of end-directedness", in [63].
137. M. J. West-Eberhard, "Adaptation: Current usages", in [81].
138. J. S. Wicken, *Evolution, Thermodynamics and Information: Extending the Darwinian Paradigm*, New York, Oxford University Press, 1987.
139. W. C. Wimsatt and J. C. Schanck, "Adaptations and the means of their avoidance", in [106].
140. A. Winfree, *When Time Breaks Down*, Princeton University Press, 1987.