In response to a direct question by Jeff Blumberg concerning Stuart Kauffman’s claim that the idea of autopoiesis can be traced back at least to Kant. I quote from an email sent by Blumberg to Autopoiesis@Thinknet.Orange.ca.us: “Dr Maturana’s answer was basically that Kauffman...
Self-organisation appears to require a sort of lifting oneself by the bootstraps without having even boots at the beginning. Self-organisation thus appears to be an oxymoron, or at least a misnomer. Autopoiesis is a self-producing process that presupposes an organised self (Maturana and Varela, 1992: 43ff). Thus self-organisation is not autopoietic in its strict sense, and inasmuch as autopoiesis or something like it is connected to autonomy, which correlates with at least one sense of “self”, the idea of self-organisation creates a quandary.

The answers I will give to questions 3 and 4 will help to explain the sense in which self-organisation is a reasonable characterisation of the processes that are so called. I will first characterise self-organisation in terms of its origin, range of application and explanatory scope, noting some open questions about the concept along the way (§2). Next I will digress to a discussion of the logic of individuation for natural systems and their properties, in which I will argue for the unique utility of a dynamically based unity relation for the determination of identity (§3). This is followed by a discussion of Bénard cell convection, often used as a Kuhnian exemplar of self-organisation, and of some other cases of self-organisation that diverge from this exemplar in important respects (§4). I will finish with an analysis of the requirements for self organisation, and discuss how these requirements entail that self-organising systems are both self-producing and self-maintaining in a clear and important sense: the very process of self organisation implies individuation of the entity formed (§5). I will conclude with some final remarks on how more developed multilayered self-organising and self-interacting systems can lead towards autonomy and a fuller sense of self control (§6).

2. Self-organisation: the basics

Self organisation occurs when the properties of a system allow it to take on a more ordered state through the dissipation of energy (production of entropy), some of which goes into the newly formed structure (Prigogine and Stengers 1984, Collier and Burch 1999, Maze and Collier, submitted). A simple example is the formation of a solar system from a primordial dust cloud, in which gravitational forces lead to collapse of the cloud and agglomeration into stars, planets, moons and asteroids. It is important to note that this agglomeration involves the dissipation of energy: in a completely conservative system, the particles would just bounce right back to their original potential energies, and there would be no formation of organisation. Agglomeration is followed by organisation into planar, nearly circular orbits. This time the dissipation originates in tidal torques. The more regular form reduces these torques, resulting in a lower energy level of the system, and harmonic entrainment of motions. In more extreme cases this process leads to resonances such as the 1:1 resonance of the Moon's orbital period with its rotation period, and the 2:3 resonance of the similar periods for Mercury. Other more subtle resonances or near resonances in the Solar System are known or postulated. Similar self-organisation can be found in macromolecules, and in lipid spheres, in which certain conformations should apply his theories in the correct domain and that he should stay out of area he knows nothing about. One thing that the good Dr did say I need clarity on. I quote Maturana: ‘...I have said that the origin of living systems occurs as a case of spontaneous organisation...notice that I speak of spontaneous, and I do so because the selfness of a system arises with the system, so a SYSTEM CANNOT ORGANISE ITSELF...’” It is worth noting that Kauffman’s (1993) models are logical, unlike the more common dynamical models used in this paper, but although proving the equivalence of the approaches is too difficult to pursue here, the basics are in (Collier 1999a).
are of lower energy. In each of these cases, the move to a lower energy conformation gives off heat (exothermic, see Wicken 1987).

Special cases of self-organisation that produce steady state systems have been described as the production of dissipative structures (Nicolis and Prigogine 1977), which are maintained in a steady state through a flow of available energy (often called exergy by engineers). Examples are Bénard convection cells and certain chemical reactions that oscillate dramatically when the appropriate heat is applied. These systems are kept in their steady state by carefully controlled boundary conditions, but more general dissipative systems occur widely in nature, such as turbulence, eddies, clouds, whirlpools and boiling in heated water. These are distinguished from dissipative structures because they are not in steady state, largely because their boundary conditions are not carefully controlled, and they are further from equilibrium. The careful control of boundary conditions in dissipative structures must not mislead one into thinking that the order arises solely from the boundary conditions, however. The system must have the appropriate internal interactions and properties to allow self-organisation in simple dissipative structures as much as in general dissipative systems that are not under such close boundary condition control. Close to equilibrium systems tend to be more orderly, and can be treated by classical methods, but far from equilibrium systems respond to applied forces by organising in the direction of the generalised coordinate of the applied force, but can respond in very unpredictable ways in other directions. In biology, Brooks and Wiley (1988) have called this the Principle of Compensatory Change. It is widespread in empirically studied biological systems, from the metabolic through the ecological levels, though not often under this name or common explanation.

The primary requirement for self-organisation, other than a source of energy to dissipate, is cohesion between the parts of the system. In the Solar System this is provided by gravity, while in Bénard cells this arises from the intermolecular interactions that produce viscosity. Photons (except when they are very dense and at high energy, permitting momentary matter to form in nonlinear electrodynamics) do not interact, and cannot be made to self organise themselves, though they can be forced to organise through interaction with matter, as in production of laser light. The other major requirement is externally applied energy in a large enough entropy gradients. This requirement can be satisfied in a wide range of ways, however, and does not add much in the way of organisation, involving the sort of order sometimes called intropy (Collier 1990a, 1999), which is the complement of entropy (the measure of disorder in a system, equal to the change of heat from absolute zero divided by the temperature; intropy is the difference between this value and the entropy value at the same temperature if all constraints were to be released). Self-organisation occurs when the system can reorganise so as to minimise entropy production (near to equilibrium systems) or reduce the entropy production in the direction of applied forces (far from equilibrium systems). Friction is a variety of entropy production, and entropy minimisation can be thought of as a sort of avoidance of friction. Self-organisation is an attempt by a system to follow the path of least resistance to energy flows through the system, and reduces the power requirements to maintain the system (Collier and Burch 1999).

It takes time for a system to respond, especially across larger extents of the system, so far from equilibrium systems in which gradients of energy and entropy are large will organise locally to some extent independently of global organisation. These leads in streams to eddies within eddies, and continual driving throughout the system maintaining these inhomogeneous conditions (coupled with the diverse effects of compensatory change) will lead to considerable diversification in strongly driven systems. Locally, systems will emerge that have lower power requirements, releasing energy for the production of even more subsystems.
These principles are not restricted to material and energy flows, but apply to information flows as well, as long as the information is understood in terms of its material basis, after the method of Brillouin (1962), originally suggested by Schrödinger (1944), according to which information is the complement of entropy (the so-called Negentropy Principle of Information, NPI) (Collier 1986, 1990a, 1990b, 1999a). For example, the information in the genome of a population can self-organise, producing a speciation event (Brooks and Wiley 1988), or the information of adaptation, which measures the correlation between hereditary features and environment, can self-organise if it is not optimal (Collier 1999b). The result is increasing information complexity and organisation, but increasing order as well, the whole system tending to use information and energy more efficiently (Collier and Burch 1999).

3. Individuation and identity: cohesion, the dividing glue

The previous section gives the basic characteristics of self-organisation, as it is usually understood, without considering the issue of whether a proper self is involved in the organisation. For that we must consider issues of individuation, unity and identity, as well as control, for the notion of a self suggests a subjectivity implying the capacity for control of the external world and perhaps the self itself. Certainly self-organisation entails a self-controlling capacity. In this section I will concentrate on the issues individuation, unity and identity, postponing the other issues until §§5 and 6.

The fundamental problem of individuation is to understand how parts of a thing can be parts of the same thing. This can be 1) through the possession of some common essence (which must be particular, perhaps a strange property for an essence, but the idea goes back to Locke, and has been taken up more recently by Putnam and Kripke), 2) a special stability (Putnam and others), 3) spatiotemporal contiguity (Wiggins 1967), or 4) a unity relation among their parts (Perry 1970). I will argue that a special unity relation called cohesion, introduced by E.O. Wiley (1981, also Brooks and Wiley 1988) to explain the coherence of species in phylogeny, and extended by myself (Collier 1988, 1999b, Collier and Muller, 1998) is the most appropriate basis for the individuation of dynamical systems and their properties. Cohesion can explain both the successes and failures of the other notions, and, in particular, being an equivalence relation, it has the right logical form for individuation.

The precise specification of cohesion is difficult because it is a matter of degree, and extends over multiple factors; the best I can do in the space available is to characterise it closely enough to make my other claims plausible. I originally introduced cohesion as the closure of the causal relations among the dynamical parts of a dynamical particular that determine its resistance to external and internal fluctuations that might disrupt its integrity. I now expand that notion to include the requirement that these relations be stronger on average within the closure than without. This determines a cohesion profile that gives the (probabilistic) conditions under which a thing will be both retain and lose its integrity, determining its boundaries under a range of conditions. We thus describe cohesion as the “dividing glue” of dynamical entities. Cohesion is an equivalence relation that partitions a set of dynamical particulars into unified and distinct entities along the lines of John Perry’s unity relation (Perry 1970).

Perry, like Locke before him, was concerned with determining what made an Fx the same Fx, especially for the case of personal identity. This problem must solve the difficulty that Ga, Hb and a and b are the same F, but it is not true that a and b are the same G or H, where F, G and H are sortals, that is, properties that individuate. For example, a boy and a man may be the same person, but they are neither the same boy nor the same man. Rather than relativize identity, Perry
suggested that strict sortals, which do ensure identity, should be understood in terms of unity, the closure of the relation that makes components parts or aspects of the same thing. The unity relation U will vary according to kind of thing. For example, Locke proposed that personal identity over time is based in memory, so if two stages of a person a and b both contain the same memories, then aUb, U in this case being determined by the closure of the relation “having a common memory”, where what constitutes memory is to be determined by empirical means together with philosophical analysis. Obviously, memory is not the unity relation for all kinds of things, and if Locke is right it is specific to people. Note further that instances of the unity relation determine equivalence classes, resulting in strict sortals, determining the identity conditions for the unified thing.

Cohesion is the unity relation for dynamical objects and their properties. Although it is quite general in application, it does not apply to natural things that are not unified by internal dynamical relations. An example of such a thing is a rainbow, which is an illusion. Other examples are apparent causal processes such as images on computer displays and theatre marquees, shotgun blasts, and, possibly, the mind, as well as more obvious examples like heaps of sand and the nomological sum of every third molecule in this room. Examples of cohesive objects are rocks, computers, human bodies and flocks of birds; probably minds, mobs and species, and possibly our universe as a whole. I leave the case of minds open not because of traditional epiphenomenalism about minds, since on dynamical views minds are not special in this respect: any apparent causal process can be mimicked by an epiphenomenal pseudoprocess (generally at some expense of parsimony). Some recent modular accounts of mind, however, suggest that separate mental processes without conscious connections are involved in producing intentional activity, despite (and contrary to) the apparent unity of mind.

Cohesion has a number of interesting properties. First, it explains the stability of cohesive entities, which is often taken to be a sign of their individuality. Cohesion is an invariant, by definition, but more important, it determines the conditions under which something will resist both externally and internally generated disruptive forces, giving the conditions for stability. Stability itself is not an equivalence relation, and does not have the correct logical form to determine individuality. Spatiotemporal contiguity has often been taken as a determinant of physical identity (Wiggins 1967). It is ensured by causal connections, as long as these are local. On the other hand, there are cases of spatiotemporal overlap of distinct systems, such as colliding galaxies, which can still be individuated by their greater self interaction than their interaction with their counterpart. These can be distinguished by spatiotemporal distinctness in the past or future, but there are chemical systems that occupy the same location but are separated by the kinetics of their processes so that they remain causally distinct despite complete spatiotemporal overlap. Such systems may be rare and contrived, but the possibility shows the superiority of cohesion as a criterion for individuation over spatiotemporal contiguity.

Secondly, the conditions underlying cohesion are just those that underlie the integrity of a cohesive thing (object or property) so that breaking the cohesion also destroys the existence of the thing. This applies to properties as much as objects. To take a simple example, the lift of a kite depends on the averaging of the impulses of individual molecules on opposite sides of the kite. This lift is not merely an averaging of the impulses (a mathematical quantity) but an actual physical average over space and time caused by the cohesion of the kite. If the kite is ripped, this cohesion is lost, the lift vanishes, and the kite plummets. The cohesion of a dynamical thing, then, determines the conditions of its existence. Cohesion, therefore, determines the Lockean real essence of a dynamical thing, what makes it what it is in fact (as opposed to what we take it to be).
A third interesting property of cohesion useful to practicing scientists as well as philosophers is its logical closure. The closure requirement can tell us what level a property or object truly belongs to by the processes and things it involves. For example, fitness is often regarded as a property of genes or organisms, but as Bob Brandon (1990, see also Collier 1998) has argued convincingly, fitness is relative to environment, and involves environmental conditions irreducibly. Thus, fitness is really a property at the ecological level, despite appearances. It is possible to abstract a gene or organism fitness, but this does damage to the real dynamical nature of fitness.

Cohesion blocks reduction of dynamical particulars because details of the microstructure of the components are irrelevant to the identity of the cohesive thing. Furthermore, in many cases, such as in gases with molecular interactions, the macroproperties like pressure are in principle not calculable because the overall motion is chaotic, though deterministic. Only the average matters. As with the kite, something must do the averaging, in this case the spatiotemporally extended cohesion of the container of the gas. In this sense, cohesion is the basis of the individuality of a dynamical object or property, since nothing smaller (or larger) is required to explain that individuality.

Many kinds of natural things are arranged in a hierarchy of existential dependence: chemical entities depend for their existence on their atomic components, biological things depend on their chemical and physical basis, psychological phenomena depend on biological processes, social occurrences depend on psychological events, and so on. Independent of questions of whether higher level disciplines are merely a special version of lower level disciplines (i.e. psychology is a specialised branch of the physical sciences), and whether higher level theories can be reduced to lower level theories, there is the question of the ontological relations among particular objects, properties and processes at different levels (i.e., the number of particular things there are).

Are higher level entities merely elaborations of the lower level processes that compose them, or do they introduce something new? Some approaches try to place emergence in novel laws, but we believe that emergent entities are prior to their laws, and that there may be entities that are emergent without new fundamental laws (Collier and Muller 1998). Phenomena that are merely elaborations are identical to some conjunctive combination of lower level processes, and are logical constructs of those processes in the sense that their identity depends on no properties not present in their lower level components. They perhaps require additional terminology, and often involve new or different means of identification and individuation, but they introduce nothing new ontologically. Such phenomena are epiphenomena. Peter Cariani (1991) calls this emergence-through-recombination, which he contrasts with emergence-through-creation. We would deny this is emergence of a new individual. Clause Emmeche and his colleagues at Copenhagen concentrate on levels of emergence. Muller and I are not convinced that there are distinct levels in all cases of emergence, given its basis in individual emergence, but our views would force agreement with Emmeche’s view that there might be “local” existence of different ontologies”. Muller and I also agree with the general thrust of their attempt to devitalise emergence through reference to nonlinear systems. But this discussion of emergence takes us some way from concerns about the self. However, in order to explicate emergence, Muller and I focussed on the individuation conditions for physical properties, objects and processes. The most important of these conditions in nonartificial contexts is self-organisation. I will describe some important cases in the next section. The central idea is that cohesion individuates dynamical systems, and gives their identity conditions, explaining their stability and spatio-temporal continuity, as well as defining their essence: without which they would not exist.
4. The Bénard cell case and some variants

Bénard convection, is one of the more intensely studied dissipative systems, both theoretically and empirically. It is therefore a useful starting point for discussing the properties of dissipative structures and developing an analogy between such structures and development and evolution. Bénard cells form when a viscous fluid is heated between two planes (or plates) in a gravitational field. The formation of the cells depends on the viscosity of the fluid, and the temperature gradient. For a given viscosity, there is a critical temperature at which fluctuations in the density of the fluid overcome the viscosity faster than they are dissipated. These fluctuations are amplified and give rise to a macroscopic circular current: dissipative structures called Bénard cells are formed. True Bénard cells have little higher order, nonlinear structure, as the conditions under which they are defined are close to equilibrium, meaning that the gradient of the relevant properties like temperature and density is greater than or equal to the size of the fluctuations.

The critical point for the transformation from conduction of heat to convection is determined by the Rayleigh number, which depends on the temperature gradient and viscosity, among other things. Chandrasekhar (1961) has what is still the best analysis of the transition. His method compared the equations of motion for convection and conduction, and set the parameters equal in order to determine the transition to the state with higher level order. There is no derivation from first principles of molecular motion, and such is probably impossible. The equations are radically nonlinear, and have no analytical solutions due to the complexity of the interactions of the molecules. (Even three gravitating bodies have no analytical solution for all cases – stability is an unpredictable property of three gravitating bodies – it is not predictable whether they form a cohesive system or not.) Prigogine’s work has shown that dissipative structures like Bénard cells minimise local entropy production, and increase the efficiency of energy throughput. Basically, the system reorganises itself to reduce the amount of friction. It follows the path of least resistance.

The properties of dissipative structures

From the Bénard cell case. We can infer several basic properties of self-organising systems:

a) Phase separation (state transition): There must be a differentiation of the system on a scale that is macroscopic relative to the scale of the components of the system.
b) Free energy source (non-equilibrium)
c) Promotion of microscopic fluctuations to macroscopic order -- formation of large-scale correlations.
d) Exportation of entropy from system.
e) System organizes itself so as to minimize local entropy production (Prigogine).
f) Energy throughout is maximized.
g) Total entropy production is maximized. (Swenson).

Point (g) is somewhat controversial. The total entropy production is increased only because the efficiency of energy throughput is increased, and it has to be dumped elsewhere. In principle it could be used to do work without further dissipation. The properties of dissipative systems are introduced in terms of a simple model: Bénard convection. In this context a number of concepts -- cohesion, emergence, epiphenomena, macrostates and microstates, hierarchy -- are
introduced. The convection cells are cohesive: their internal interactions are stronger than their interactions with other cells, thus we can talk of the individuation of the cells. The convection is emergent, since it is maintained by cohesion, and is not merely an epiphenomenon, as the circular motion of widely spread photons in a high gravitational field would be – in that case the photons do not interact, and there is no self-organisation of their motion – it is imposed by external conditions alone. This brings us to another very important point about self-organisation: it requires the properties of the organising parts to be of the right form to allow organisation to form. Without this, the organisation is imposed entirely by external forces. Self-organisation leads to macro properties that don’t exist at the microscopic scale. They are properties that extend over space and time (non-local). This is typical of cohesion, as in the kite example above. Entirely local physics is not possible: recall that Chandrasekhar assumed, but did not deduce, the convective state.

Species and Development as Products of Self Organization?

When the Bénard cell analogy introduced in the beginning of this chapter is reexamined and compared to other individuating processes involving self-organisation a number of disanalogies appear. The process of speciation seems consistent with the properties of dissipative structures under certain conditions. Living systems may have additional properties that are not encompassed by simple models of dissipative systems (this discussion is based on Collier et al submitted).

Fluctuations in gene transfer due to accident, mate choice or external factors such as preferences of different genotypes for different food sources will segregate the population genome. As the genetic information common to the population decreases overall, there will remain subgroups in which the common information (within that subgroup) is the same as it was previously for the whole group. These groups will now appear to have a degree of cohesion formerly attributed to the group as a whole. That entity which was the group as a whole will still be attributed a certain cohesion, but at a different scale. What has just been described is the process of population differentiation leading to speciation. It is also analogous to the description of Bénard cell formation, with variants implying fluctuations, and common information due to vertical or horizontal constraint implying cohesion (which will act analogously to viscosity). However, there are a series of disanalogies that should be noted.

1) In convection there is no differentiation.
2) In convection the applied force acts uniformly.
3) Hence the analogy is imperfect because convection occurs in a uniform fluid, and there is no sorting. The disanalogies appears specifically between gravity and differentiating forces. Gravity acts on all molecules equally, whereas differentiating forces are presumed to act differentially on different phenotypes. Thus Bénard convection represents a starting point for thinking about speciation as a form of self organization, but the endpoint is obviously more complex than the model implies or encompasses.

As simple model of speciation of this form involves three genotypes with two alleles at a common locus, with TT producing tall organisms, tt producing short organisms, and Tt and tT producing intermediate height organisms. One way speciation can occur is if the TTs develop a preference for each other. The Tt and tT forms will disappear, and if the TT and tt forms cannot mate, speciation occurs. Alternatively, and environmental change might lead to trees invading bushland, and the TTs tend to feed on the trees, mating preferentially with those nearby, leading
again to the disappearance of the Tt and tT forms, and speciation. In small populations, chance fluctuations can lead to the same effect. How often this sort of sympatric speciation occurs is an empirical question that has not been well studied. It is also unclear how strong the effect is compared to selective factors. Nonetheless, assortative mating is known in nature, and it should be studied to see if there are the sort of non-equilibrium causes described. Note that no differences in fitness (reproductive success) are required for this model of speciation.

Similar self-organisation may be involved in development (Collier et al submitted). Meristems are equipotent plant cells that differentiate into diverse forms. Usually this is thought to be under genetic control, but this is just a hypothesis. Another possible mechanism is differentiation through fluctuations. This may explain why genetically similar plants in similar environments show different structures (Maze et al 1970, Maze et al 1971, Maze et al 1986, Maze et al 1987, Maze et al 1990, Maze and Vyse 1993, Maze and Collier, submitted). Such cases show the same sort of disanalogies with convection cell formation, and solar system organisation. A plausible hypothesis is that internal factors are far more important in determining the differentiation than external factors, even given the chance factor required in self-organisation. It is this imbalance of determining factors, if anything, that puts the “self” in self-organisation. Note that this sort of process can produce individuality even in a situation where there is a common genome and common environment. The differentiation further allows the mechanisms of self-organising speciation to work.

More complex forms of organisation are found in autonomous systems, discussed further in §6. They are characterised by large scale correlations in the information in the system, a sign (and product of) logical depth (Bennett 1985), which is a measure of time required to compute the surface structure from the most compressed representation of something. Organisation can appear through design, in which case it is derived, from the influence of external forces causing selection, or through self-organisation. One characteristic of creative autonomous systems is their capacity to use the last method to create new organisation (Collier 1999c). This is perhaps the clearest case of true self-organisation, since the system conditions themselves govern its further organisation without further influence. Maturana does not seem to have considered this possibility, but he should consider it a case of self-organisation even in the strictest sense.

5. Requirements for self-organisation

Self-organisation requires an entropy gradient in which the entropy can be given a physical interpretation compatible with statistical mechanics, though it does not need to be the usual entropies of physical or chemical systems, but can include information systems or multicomponent genetic systems like those in Kauffman’s models. The entropy gradient must exist across the boundaries of the parts of the system that self-organise. In physical and chemical cases, this requires a source of available energy, or exergy, that can be degraded such that the total entropy production can be negative within the system, producing new order and organisation, but without violating the second law. Because of the connections between energy and physical interpretations of information (Brillouin 1962, Collier 1986, Collier 1999a), an exergy source is also required for informational self-organisation. I assume that a similar requirement is required by Kauffman’s Boolean systems, but a proof would be nice. This is the most fundamental requirement for self-organisation. It is external to the system, but it is very nonspecific, and contains little required organisational information – it is not a source of organisation itself. This provides part of the answer to question (4): lacking required organisation, exergy supplies do not modulate or control.

The second requirement for self-organisation is that the components of the system interact with each other. Weakly distributed photons cannot self-organise, nor can the ideal molecules in
an ideal gas. Something like the viscosity of fluids, sexual interaction in species, or metabolic interactions in development are required. This fundamental requirement is needed to permit the large scale correlations found in organisation (unless they are accidental, which is improbable). Note that this condition is one of the components of the system and their relations to each other, and need be to nothing external. This goes some way to answering question 3): the conditions that contain the information (discrimination or differentiation) required for self-organisation are inherent to the system components themselves. This is indicated in the differing fluctuations in a non-convecting fluid, the genetic differentiation of the model population described above, and the differing potencies of meristems in botanical systems.

A third requirement for self-organisation is a critical point between different phases of the system, beyond which entropy production is decreased. There may be only, or two or more of these, which are attractors for a system that is driven strongly enough by an entropy gradient. Which attractor is occupied is largely a matter of chance created by fluctuations within the system. Again, this significant condition is system internal, and is not modulated by any significant external conditions.

The result of self-organisation is increased stability of the system, a slowing of energy or information flow, along with an increase in the efficiency of this flow. The system also individuates and in more complex cases can organise with considerable logical depth. Complexly organised systems, however, generally are formed by an iterative process of self-organisation and selection of less successful forms.

Since the external gradient need contain little organisation or information of other forms except entropy/exergy, which is statistical in nature, and undifferentiated relative to system organisation in self-organising cases, the process is not externally modulated. Internal conditions like component interactions and the existence of attractors are the source of the possibility of self organisation. The process is inevitable if the entropy gradient is large enough, but in systems with two or more attractors, the path taken is not predictable. Nonetheless, it is internal conditions that lead to the individuation that we call “self”. The special properties of the fluctuations that make self-organisation possible are the proto-selves that organise into individuated selves by the process of self-organisation, it is not unreasonable to say that they organise themselves. There is still room to disagree with this usage, but there is less so for the special case of autonomous systems, which I will consider next.

6. Individuality and autonomy
A system is autonomous if it uses its own information to modify itself and its environment to enhance its survival, responding to both environmental and internal stimuli to modify its basic functions to increase its viability. A major constraint on the survival of an artifact is that it serves its designed purpose: A household robot that makes messes will not last long. Similarly an organism will not last long if its functioning does not contribute well to its autonomy; it will be selected against by natural selection. This inverts, or perhaps more accurately, complements, the currently popular etiological accounts of function, according to which a function’s purpose is that for which it is selected (Wright 1973, Millikan 1989, Neander 1991, but see Christensen 1996, Christensen and Hooker 1997, Foss 1994, Christensen et al submitted). The basic idea of the etiological view is that a property P is selected because it does F, and because it does F the organism that possesses P is selected. Instead, on the autonomy view, the autonomy found in the organisation of some things (especially organisms and lineages of organisms) which includes F among its functions contributing to autonomy, sustains their viability and likelihood of being selected. Thus the basic function of P is to contribute to autonomy, which in turn makes the
organism more viable than it would be without \( P \) and all other things being equal. Selection is a result of functionality on this account, not its cause, though it is the cause of its retention. The self-referential and open character of autonomy requires that an autonomous system be flexible, open to signals, capable of self-modification of any of its anticipatory functions, and capable of evaluating such modifications. A useful, but not necessary characteristic is the second order capacity to anticipatively self-modify, permitting self-driven adaptability. This higher order property requires some sort of self representation and recognition of what could contribute to autonomy, since it is not directly subject to natural selection to weed out unsuccessful modifications. In any case, the self-modifications would not have a genetic basis, so selection would not preserve them unless the modifications somehow became genetically fixed. Internal (vicarious) selection may play an important role, perhaps guided by external positive and negative stimuli, but the modifications are basically self-guided. (I ignore the possibility of an organism or device with such a propensity for pathological self-modification that it has little chance for survival; I assume that vicarious selection will eliminate hair-brained schemes, see (Bickhard and Terveen 1995).

Naturally autonomous systems have a dynamical (causal) cohesion (Collier 1988, Collier 1999c, Christensen et al, submitted) that is actively maintained by internal and external processes of various kinds that are controlled by their internal information, i.e. they are substantially dynamically self-maintaining. Cohesion is the closure of unity relation among parts of a dynamical system comprised by the closure of their dynamical (including functional) processes that maintain system integrity in the face of external and internal fluctuations. Cohesion individuates systems by the internal cohesion being generally stronger than internal fluctuations and external insults. For example, in a predator both active and passive cohesion (e.g. largely active food searching behaviour followed by eating and metabolism vs the relatively passive structure of the bones) help to maintain system integrity while at the same time serving to differentiate one system from another; no two predators interact physiologically or metabolically with each other more than they interact in these ways with themselves.

Autonomous systems have many functional properties that preserve system properties through cycles of interaction, both internally and with the environment. These cycles are typically complex and self-reinforcing. Process closure concerns the fact that an overall process must achieve self-reinforcement by supporting system viability, and hence the continuing system capacity to carry out that process. If the system is to achieve overall process closure the elements of the system must interact with each other and with the environment in particular, circumscribed ways. This is interaction closure. It is essential to self-regulation, and distinguishes autonomous systems from other cohesive systems like rocks that maintain their integrity merely through strong bonds that tend to isolate them from other systems, and from systems like gases, and liquids that are more open than solids, but do not have any closure of environmental interaction required for self-regulation; they remain independent only at the whim of environmental contingencies. Although open-ended interaction with the environment makes autonomy a property at the ecological level, in the sense that the closure conditions for its definition make essential reference to the environment, autonomy “belongs” to the autonomous individual in the sense that what makes the difference to autonomy (the information controlling cohesion) lies in the individual.

Autonomous system processes will in general interact with many other such processes, e.g. eating and digesting, etc., support not just hunting capacity that can lead to further eating and digestion, but every other system capacity as well. To maintain themselves autonomous systems must display a corresponding internal coherency of processes; namely the processes must interrelate flexibly so as to preserve the whole organised complexity that underwrites control of
that very responsiveness and adaptability. Their functional properties must be so integrated that they can maintain an active independence. Unlike all other kinds of systems, autonomous systems are dominated by these global functional constraints.

Autonomy is multidimensional and varies in degree. If the dimensions are distinct enough, we can talk of kinds of autonomy, such as material autonomy, psychological autonomy, social autonomy and informational autonomy. These kinds of autonomy are at different levels, so autonomy is also relative to level. Something that might not be autonomous at the most fundamental physical level, under the extreme conditions found in physics, but might be autonomous biologically in the less intense environments of organisms, and minds, for example, might be autonomous in terms of information content, even though they depend on their biological embodiment. In each case, a higher level autonomy will require the existence of an underlying autonomous system, and a kind or level of autonomy will usually contribute to the autonomy of its constituting and embedding level; however, levels and kinds of autonomy can compete just like autonomous individuals compete at the same level. Autonomy may be largely in one dimension or interdependent range of dimensions, despite large dependencies of other kinds. It might be argued, for example, that minds, although highly dependent on bodies materialy, are informationally quite autonomous, as would be other autopoietic entities (Maturana and Varela, 1972). The use of the body by the mind to maintain itself, it might be argued, is analogous to the use of the environment by the body to maintain itself, creating not only an informational independence but arguably making it self-sustaining as well. This maintenance can conflict with bodily function in extreme cases, such as when a hero satisfies his self conception of his personal integrity, and sacrifices his life for unrelated others. If representations can be autonomous, their autonomy will be of this informational kind: they must actively use their own information to maintain their own informational structure and reproduction. Fundamentally, though, they must function to preserve the autonomy of those who have the representations; a wholly “selfish” meme or crazy idea or ideology would soon disappear, if only because its possessors would not survive.

An autonomous device in itself is not especially useful; on the contrary, its behaviour on training may be “perverse”, since it will respond to enhance its autonomy (and often not well, at that), not the design goals of its training. This can be seen in training natural systems like animals, children and ecologies, in which our best efforts to control the system often leads to the opposite effect. The same frustrations should be expected in training an autonomous robot. Part of the problem is that merely autonomous systems are likely to have minimal anticipatory capacity unless they have sophisticated representational capacity. This is true, for example of young children and ecosystems, which lack the representational capacity to convert verbal ideas or demonstrations into internally governed practices. In such cases, modification of the system behaviour requires the simultaneous formation and integration of the required structures and dynamical relations to achieve the desired end. This requires both patience and a good understanding of the system. The trainer has an idea what he or she wants the trainee to learn, but this may not by easy to make compatible with the requirements of autonomy of the trainee. For example, forcing children to eat is usually counterproductive, creating strife at meals. Alternatively, a hungry child will eat some food it does not like, and may come to like the food if it becomes associated and integrated with pleasant experiences. Forcing eating just associates food with unpleasant experiences, and further integrates resistance to eating into the child’s autonomy. The point is that the child can maintain its autonomy with a choice, by refusing to eat and thereby maintaining control of the situation, or control can be removed as an issue, and the child can learn on its own proper eating habits.
With the varieties of maintenance of autonomy in mind, the forcing route is likely to be counterproductive.

Autonomy is always self generating, or an autonomous system would not be able to maintain itself. Furthermore, autonomous systems are best formed spontaneously through the integration (through functional and structural cohesion) of their prior properties, altering those properties so that they are constrained by the newly formed cohesion (in other words, they are emergent). They could, in principle, be designed, but the risk is that constraints will be built in that produce a device limited by conscious or unconscious design constraints that prevent spontaneous self-organisation, both through restrictions on interactions with the environment, and on internal reorganisation to respond to unexpected signals and especially of unexpected signal types (Cariani, 1991; 1993). The best way to produce an autonomous device is to let it grow under the right conditions. Thus the problem of devising a self-modifying anticipatory device that can develop modifications even to what it can recognise and control is more analogous in some respects to horticulture than to mechanical manufacture.

As mentioned above, Maturana’s objections to the notion of self-organisation apply most directly to autonomous systems of the sort I have described. This is ironic, since Maturana and Varela consider autopoiesis to be the basis of individuality and autonomy. I believe the problem is that they have not properly considered the role of self-organisation in autonomous systems, or for that matter, taken self-organisation seriously at all. At best, an autopoietic system can be self-governing, and rearrange its organisation, but it cannot produce new organisation. This is more akin to a mechanistic device than biological systems or human minds.

7. Conclusion

Self-organisation requires an entropy gradient that is external. But this need contain no further organisation. The important information for self-organisation is found within the system itself, and is not modulated by external organisation. Self-organisation produces individuation, and thus new “selves” using the properties of the component parts under the right conditions. For this reason, the major significance of the information internal to the process of self-organisation, it is not unreasonable to say that it is the self that organises. This process is most pronounced in creative autonomous systems, which ironically are exactly the ones that Maturana and Varela though autopoiesis was sufficient to explain. They were wrong.

Acknowledgments

Much of the work in this paper is based on collaboration with the Newcastle Dynamic Organised Complex Adaptive Systems Group, especially C.A. Hooker, W.D. Christensen and S.J. Muller.

References

Cariani (1993) gives an example of a physical device invented by Pask in the 1950's that could modify its electrical and electrochemical ferrous sulphate substrate to distinguish tones and magnetic fields. These capacities formed spontaneously through changes in the malleable ferrous sulphate substrate under exposure to appropriate stimuli, rather than being designed in from the beginning. Unfortunately, this work was not pursued further, perhaps due to the dominant computational model of mind.


