Chapter II

Autonomy

W.D. Christensen, J.D. Collier, C.A. Hooker

Contents

1. Introduction
   1.1 Global functional constraints
   1.2 Directed processes
   1.3 Autonomy, directedness and normative perspective
   1.4 Modulatory processes and hierarchical organisation
   1.5 Interaction and the locus of directive constraint

2. Anticipation
   2.1 Anticipative Closure: General Specifications
   2.2 Anticipative Closure: Detailed Specifications

3. Modelling autonomy dynamically

4. Conclusion

---

1This chapter is adapted from ‘Autonomy, Adaptiveness, Anticipation: Towards autonomy-theoretic foundations for life and intelligence in complex adaptive self-organising systems’, in preparation. A briefer account of autonomy may be found in Christensen and Hooker 1998b.
II: Autonomy

1. Introduction

The concept of autonomy is designed to capture the general organisational nature of living systems. If intelligence is firmly rooted in natural life capacities, as was suggested in chapter I, then a foundational theory of this type is necessary in order to properly understand the nature and emergence of intelligent systems. Living systems are a particular kind of cohesive system, where a cohesive system is one in which there are dynamical bonds amongst the elements of the system which individuate the system from its environment (see Collier 1988, Christensen, Collier and Hooker 1997a). These bonds fall into different organisational kinds — some cohesive systems are based on stable structural relationships which cause the components of the system to bind together statically (e.g. rocks), whilst others are based on process relationships which continuously re-create the system (e.g. cells). The latter are what is here called autonomous systems, systems whose integrity arises from self-generating, self-reinforcing processes.

To gain an intuitive feel for the systemic distinctions being made compare a gas, a rock, and a living cell. A gas has no internal cohesion, it takes whatever shape and condition its containing environment imposes and will simply disperse if not externally constrained. By contrast, a rock possesses internal bonds which constrain the behaviour of its elements in such a way that the rock behaves dynamically as an integral whole. The notable organisational features of these cohesive bonds are passivity, rigidity and localisation. The bonds are passive and rigid in that they are stable deep energy well interactions which constrain the constituent molecules to spatial positions within a crystal lattice. The bonds are localised in the sense that the strength of the forces which bind a molecule within the crystal lattice depend only on the connections with adjacent molecules. This localisation means that there are no essential constraints on where the boundaries of the rock must lie — if it is split the particularity of the rock’s identity is disrupted, but the result is two smaller rocks with exactly the same type of cohesion properties as the original.
II: Autonomy

A living cell is similar to a rock inasmuch as it possesses cohesive bonds which cause it to behave as an integrated whole, however organisationally it is very different to the rock. In particular, the cohesion of a cell is active, flexible and holistic. The chemical bonds of a cell are formed by shallow energy well interactions; they have short time scales relative to the life of the cell and must be continually actively remade with the assistance of external energy fluxes. This continuous activity makes the cell vulnerable to disruption but also gives the cell flexibility since the interactions can vary according to circumstances by responding sensitively to system and environmental changes. The cohesion of a cell is holistic because the forces which bind its parts depend on globally organised interactions. That is, local interactions must form functional processes that interact at the global level of the cell to reproduce the conditions necessary for the cell’s survival. As a result of this holistic organisation, cutting a cell in two usually does not produce two new cells (in contrast with the rock) because the processes which maintain the cell are disrupted.

Autonomous systems are cohesive systems whose organisation is of the same general type as the cell. That is, autonomous systems are actively self-maintaining systems constituted in complex processes that are sustained by open cycles of interaction, internally and with the environment. This means that the cohesion conditions of autonomous systems: (1) tend to rely on relatively shallow energy wells, (2) are nonstationary, with the dynamical conditions underpinning cohesion being characteristically oscillatory or chaotic in nature, (3) rely essentially on self-generated dynamical conditions to direct resources into cohesion-maintaining processes and (4) achieve dynamical self-generation through the possession of internal organisation which performs work to direct energy fluxes from the environment into the constitutive processes of the system. Condition 3 identifies autonomous systems as being types of positive feedback systems, typically with stabilising negative feedback as well. Condition 4 specifies that the locus of control in autonomous systems is substantially internal (but see §1.5 below). Condition 4 distinguishes genuinely autonomous systems from other kinds of phase-separated positive feedback systems; in principle from systems, such as candle flames, which exercise no self-reproductive
II: Autonomy

control (whence the determination of system features lie outside of whatever organisation there is), and in degree from systems, such as viral parasites, where relatively little (but not no) control of reproduction resides within the system.

Examples of autonomous systems include the following:

- **Molecular catalytic bi-cycles**: Mutually catalytic molecules which form a self-sustaining bi-cycle system (Rebek Jr. 1994). This is the minimal case of a self-generated process-based system. The directive organisation (see §1.2 below) which generates the process-patterning lies in the macro-molecular conformation which catalyses the formation of the mirror template molecule from the substrate material.

- **Organisms**: The paradigm examples of autonomous systems are uni- and multi-cellular organisms, where cell and skin membranes differentiate internal and external environments; metabolic systems maintain critical physiological parameters for system functioning (pH level, temperature, stored energy in forms such as ATP); and, for the more deeply complex multi-cellular organisms, an immune system destroys harmful invaders while a sensori-motor/cognitive system regulates environmental interaction, seeking out critical resources (food, water, shelter, mate) and avoiding danger (poisons, predators etc.).

- **Species**: The autonomy of a species lies in the way the population is maintained over time through evolutionary adaptation. Mutation and ontogeny (the feedforward aspects of evolution) allow a species to explore its genomic and phenotypic configurational possibilities. Natural selection (the feedback aspect of evolution) eliminates unfit phenotypes (and their corresponding genomes). The net effect is that the species forms an autonomous process maintained by a feedback cycle between the population and the environment as fit genomic-ontogenetic configurations proliferate and the species as a whole explores its organisational possibilities, evolving to its accessible, environmentally successful organisational forms. Together, genetic and ontogenetic exploration and negative feedback result in the maintenance of a
II: Autonomy

stable macro cycle.²

· Colonies: Human cities, e.g., actively import the resources from around them essential to maintaining their functions (water, foods, fuels, materials, information ...) and distribute these through complex transport pathways so as to preserve their functional integrity, they have many processes which respond to deficiencies in supply and other internal threats to coherence, from private procurement in markets and charity work to public regulation of procurement (e.g. for water), internal restitution (e.g. educational production) and regulation (e.g. policing, cf. immune function). It is this complex form of organisation which accounts for the fact that cities are simultaneously both very resilient in some respects and highly fragile in others.

· Ecosystems: interspecies relations organise and complexify matter/energy throughput, stabilising far-from-equilibrium structures.

Thus, although all autonomous systems share characteristic features, there is also considerable organisational variety.

Autonomous systems are cohesive self-maintenant systems, but it would be a mistake to interpret the term ‘autonomy’ as implying complete independence from the environment. Indeed, there are at least three clear ways in which autonomous systems are not independent of their environment: (1) as dynamically open systems, autonomous systems are coupled to their environments by nonlinear interactions and hence cannot be analytically decomposed as a linear sum of system plus environment, (2) as far-from-equilibrium dissipative systems, autonomous systems require energy input from the environment, (3) as adaptive systems, the functional organisation of autonomous systems must be characterised in relation to at least some of the determinable features of the environment, indeed to the extent that they have organisational depth³ because they rely on environmental order, their depth must be

---

²This is ultimately the principled dynamical basis for any claim that species are themselves individuals, see e.g. Hull 1988.
³See below for a technical discussion or organisational depth. The intuitive idea is of organisational richness.
II: Autonomy

classified in terms of that order.

So, rather than involving complete independence from the environment, autonomy as it is being theorised here involves a certain kind of organisational asymmetry between the system and environment, namely that the directive organisation which induces the pattern-formation of energy flows from the environmental milieu into system-constitutive processes is substantially endogenous to the system itself. Although aspects of the environment participate in the overall process-cycle in and by which the system is constituted, they require the system’s directed processes to become channelled into system-distinctive processes. For example, a particular bird species may depend on the presence of twigs to make nests, but it does not depend on any particular twigs since the birds can choose from what is available and twigs of themselves have no tendency to play a role in bird-creating processes unless co-opted by birds. So whilst autonomous systems depend on dynamical and organisational features of the environment (twigs, for example, are organised), they are distinctively characterised by internal directive organisation and consequent pattern creation capacity upon which their existence depends.

Not all dissipative processes constitute autonomous systems in the sense outlined above because substantial parts of the directive organisation critical to their existence lies outside the system. Many process-based systems, such as Bénard cells, are wholly determined by the presence of external energy fluxes; these systems are driven by their environment. Other systems contain some of the directive organisation necessary for their existence endogenously, however this directive organisation is incomplete and they rely essentially on external sources of organisation to structure the processes essential for their cohesion. Viruses, for example, rely on the genetic machinery of the host cell to reproduce. Directively incomplete systems must rely on relations with other systems to achieve self-generative closure; the two principle forms these relations can take are parasitism and symbiosis. Parasitic systems depend on a distinct autonomous supersystem (i.e., one that is independently autonomous) for self-generation. Symbiosis involves coupled systems that rely on mutual interaction
II: Autonomy

for self-generation. This cooperative coupling forms an autonomous supersystem whose boundaries are determined by ultimate process closure (the full set of processes required to form a self-generative system), and which cannot exist without these interactions. Symbiotic autonomous supersystems is the basic I-C model for understanding the emergence of multicellular organisms, some features of ecosystems, and culture (see chapter III §7). With respect to the latter, it is interesting and important to note that humans are an essentially culturally symbiotic species — the isolated individual possessed of no cultural learning is incapable of survival.

Thus, while heterogeneous, autonomous systems form a natural kind because underlying their surface variation these systems share a common process/constraint topology which is fundamental to their existence. What characterises all these organisational types in common is the centrality of holistic organisational constraints on their cohesion. Indeed, it is presumed that a substantial amount of the variation displayed by extant autonomous systems can be accounted for in terms of their varying locations in the space created by the dimensionality of their fundamental process topology (less formally, by their differing in the many different respects which characterise their autonomous process organisation). Next, some of this dimensionality is explored.

1.1 *Global functional constraints*

Unlike all other natural kinds of systems, the identity conditions of autonomous systems are dominated by global functional constraints. This feature was illustrated in the contrast between rocks and cells discussed above. As noted then, the molecular cohesion conditions of a rock are local, they place no particular constraints on where its boundaries must be — if it is split the particularity of the rock’s identity is disrupted, but the result is two rocks with exactly the same type of cohesion properties as the original. The cohesion properties of an autonomous system, on the other hand, typically cannot tolerate disruption to the system’s particular identity. If you cut a cell in two you typically do not get two new cells, because the fundamental
II: Autonomy

processes which produce cell-type cohesion have been disrupted. Autonomous systems are constituted in and by their global functional organisation: In contrast to a rock, the identity of a cell constrains its fundamental cohesion conditions. Nonetheless, as also noted, these organisational properties may be realised by diverse specific interaction schemes which vary enormously in their low level and intermediate properties.

Possessing a global functional organisation means that the constitutive processes of autonomous systems are interrelated so as to satisfy certain global properties, principle amongst which are two related properties, labelled process closure and interaction closure. Understanding these global functional constraints on autonomous systems involves looking at the diachronic structure of these systems: to properly specify the identity conditions of an autonomous system it is necessary to incorporate, not just the synchronic relationships amongst the elements of the system, but also the characteristic time frame over which the system’s processes unfold and the relationships between stages of those processes. Process closure and interaction closure are the two principle dimensions of the anticipative closure conditions of autonomous systems discussed in §2 below and play a key role in adaptiveness (discussed in chapter III).

The process closure constraint concerns the fact that the overall process scheme of an autonomous system must be self-generating in this sense: the overall state of an autonomous system will tend to create the conditions under which the system retains its cohesion, and hence also its self-generating capacity, in the future. Frequently, this is achieved through a cyclic process, e.g.: A → B → C → A, where A, B & C are process stages. Cycles involve quite concrete process closure inasmuch as the initial stage of the process must be a product of the final stage, e.g. the birth → maturation → sex → birth cycle of an organism lineage. However not all processes in autonomous systems are cyclic, and closure can take the form of a high ordered rather than concrete condition, hence autonomous systems can (and often do) display quasi-

4Roughly, globally specified. See below for a more precise specification of organisational order.
II: Autonomy

periodic and chaotic dynamics (see below). Typically autonomous systems will be constituted by many internested processes which inter-mesh in complex ways to form an overall system-environment cycle of interaction which is self-generating for extended periods of time. With respect to this diachronic process organisation, process stages are essentially feedforward actions in the sense that they must result in certain types of outcome if the system is to achieve overall process closure. We can abstract away from the details of the system’s processes and identify process closure as the general property of sustained positive feedback running from the system’s interactions back to the system’s processes. This has the form of an open loop: system → action → environment → system. Whether a system achieves process closure is a context-dependent property, which is why autonomy can only be specified relative to an environmental context.

Because natural autonomous systems are dynamically constituted, achieving process closure involves achieving a closely related form of success — interaction closure. The elements of a system interact with each other and with the environment in circumscribed ways; if the system is to achieve overall process closure, these ways must so interrelate that they dynamically generate that closure. In the cyclic example above, it must be the case that A actually produces B and, in some suitable range of internal conditions obtaining in the organism, produces it in sufficient quantity and sufficiently resiliently, and so on for the other interactants. Again, the multifarious biochemical reactions that constitute hunting activity must actually so interact with the environment that a sufficient supply of organic material results and that material must be of a kind that can actually be ingested and digested in the gut and with a sufficiently nutritive result that subsequent hunting activity (among others) is biochemically sustainable. None of these conditions is trivial. Most autonomous systems are able to achieve overall closure because they possess high levels of internal organisation in the form of an organised chemical metabolism. They extract resources from the environment and process them in structured ways to produce the substrates for the system processes themselves. Though they permit many diverse viable functional organisations, interactive closure provides the underlying complex
II: Autonomy

of dynamical constraints on all process closure.

Because of their global functional organisation and feedforward process character, with concomitant process and interaction closure constraints, autonomous systems are teleological. Indeed, they are the most fundamental form of teleological system. Because of their perpetual dependence on positive reinforcement from the environment, the existence of autonomous systems hinges on their capacity to satisfy the afore-mentioned process and interaction closure conditions. These conditions constitute a basic emergent normative framework for system behaviour, discussed further below.

1.2 Directed processes

In characterising organised processes it is necessary to distinguish between directedness, regulation and control. Directedness involves constraints which structure energy flows, e.g. in the way that the banks of a river structure the flow of the water. Regulation involves constraints which preserve some form of dynamical stability. Control involves the maintenance of a dynamical state through feedback indexed to a set point. Directedness is the most general form of organised process relation involved in autonomy, though it should be noted that in the context of autonomy a directed process is subject to normative closure constraints (as a river is not).

All three types of structuring process play a role in the adaptively organised processes of autonomous systems, however their material realisations and mutual interrelationships can be extremely subtle and complex. This complexity arises because of the highly multiplexed nature of living processes, whereby multiple functional roles – each with a distinctive form of closure – can be performed by a

---

5Two dynamical variables, say A and B, may be dynamically related by interaction, regulation or control. A dynamically interacts with B if A dynamically influences B, A regulates B if A's dynamical interaction with B induces dynamic stability in B, A controls B if A regulates B in dynamical relation to the difference between B's state and that of some reference state R. These relations may (though need not) also obtain between B and A and, generalising, they may obtain between any number of dynamical variables.
II: Autonomy

single dynamical process. Most generally, autonomous systems need to preserve their autonomy, but this will rarely occur through literal control (i.e., through feedback control). Instead it will usually be achieved through the imposition of directive constraints which in fact generate the closure conditions for autonomy without directly controlling for them.

Much of the complexity associated with understanding teleology and normative function in biological systems stems from the complex relations between directed processes and closure conditions. There can be multiple closure conditions associated with a directed process, some of which are operationally explicit and some of which – typically the deeper autonomy-related functions – are implicit. For example, many organisms directly control for hunger satiation, but do not directly control for nutritional intake. The difference is revealed in the fact that organisms, including humans, will consume pseudo-foods which look and taste like food and reduce hunger pangs but have no nutritional content. Only humans ever explicitly control for nutrition, and they only do this by selecting foods on the basis of nutritional analysis. Thus, by controlling for hunger satiation organisms (mostly) regulate nutritional intake, and also autonomy. Other process aspects, such as specific foods consumed, can be directly structured without being regulated. The organism might simply consume whatever things it ‘recognises’ as food which it comes across, without maintaining any balance or quantity of food types. If so, it directly structures its food intake (via its food recognition processes), without regulating this intake.

These distinctions are important for understanding adaptive interaction since the way a process is organised to achieve its closure conditions has an important impact on the system’s interaction capacity, affecting its openness, capacity to respond to variation, and capacity to learn. For instance, organisms which do not regulate for balanced food intake will not respond to an environmentally induced variation in food types which affects nutrition by adapting their food collecting behaviour, whereas an organism that regulates for balanced food intake can vary types of food intake to maintain nutrition, and hence respond adaptively to changes in food availability.
And, as will be argued in chapters III and V, whilst learning processes must be regulated or controlled in certain respects in order to be adaptive, they must also incorporate directly structured uncontrolled plasticity in order to be open.

A further important aspect of directive constraints is the extent of hierarchical or quasi-hierarchical structure they possess. This is determined by the degree of modulatory asymmetry which occurs in the system, or, to put it another way, by the order of the directive constraints. Technically in information complexity theory order is specified in terms of correlation, where the order of a correlation is the number of entities which have to be specified to specify the correlation. In this latter sense living systems and most other autonomous systems, and all of the cognitively complex processes later discussed, have many very high order constraints, processes etc. because their orderedness has to be intrinsically global, it has to continuously reproduce the process interrelationships which generate the system’s integrity. (For a more detailed discussion of order in living systems see Collier and Hooker 1998.)

Intuitively, to accomplish this they must also be, and are, highly organised. One important measure of the sophistication of organisation is organisational depth, which roughly refers to the complexity of the correlations within the system, more technically to the redundancy order of the system in the above sense: A system with high order correlations shows corresponding subtle high order redundancies. This can also be expressed as having great logical depth, measured by the number of expansions between the most compressed description of a system (using its correlations to compress description) and its most detailed explicit description. From a dynamical perspective deeply organised systems display subtle, globally coherent organised processes.

We will also use related notions of regulatory and control order, where in each case the order of system correlations among components which is required to realise the

---

6See Bennett 1985. The information theoretic concept of order is an important guide to understanding organisation in the systems of concern here, but is insufficient of itself essentially because it is not dynamical (Collier and Hooker 1998); however it is not relevant to pursue these issues here.
II: Autonomy

relation determines the order of that relation. In this way local relationship orderings can also be distinguished, e.g. intuitively \{C controls (B controls A)\} is a higher order relation than the *chained* relation \{(C controls B) and (B controls A)\}, which is in turn is a higher order relation than \(A \text{ controls } B\). Control ordering can also vary, e.g. where at time \(t_1\) \(A\) controls \(B\) while at \(t_2\) \(B\) controls \(A\), or where \(C\) switches which controls which, or even switches their relation to regulation or influence. To satisfy their organisational requirements, living systems are replete with context-sensitive regulatory and control relations, many of high order.

1.3 *Autonomy, directedness and normative perspective*

This excursion into the metaphysics of organisation, directedness and control prepares the ground for understanding the degrees and interrelationships between endogenous and environmental directedness in autonomous systems; roughly, the *endogenous locus of control* identified above as a characteristic of autonomy.

Autonomous systems possess directive organisation which structures the system’s process flow in ways which satisfy the conditions sustaining the system’s integrity. In other words, directed processes in autonomous systems must satisfy normative constraints in the sense that they must keep the system’s state within some delimited range of parameter values, on pain of the system otherwise ceasing to be that kind of system, of losing its identity-constituting raison d’être. In that sense the system ought to maintain its state within that delimited range of parameter values.

This characterisation makes it appear that autonomy is essentially a control function, but this is not the case. The classical engineering model of control is of a system \(S\) coupled with a controller \(C\) which modulates \(S\)’s dynamical behaviour with respect to a set point such that \(S+C\) realises \(R\) — the reference output, or goal, of the system.\(^7\)

The set point, and the modulatory actions of the controller, determine the evaluative framework against which actual system performance is measured, in effect

---

\(^7\)The definition of control given above preserves this sense of control as an organised process, but recognises a richer set of organisational relations.
II: Autonomy

establishing a *normative perspective* for the system. However, as observed, the directed processes of autonomous systems may achieve their closure conditions without directly controlling for them. They possess a normative perspective, but not one which depends on literal control.

Understanding the origins and nature of normativity in natural systems is no simple matter, and it is a fairly widespread view in philosophy that there is no causal basis for norms. Naturalistic approaches to normativity typically hold that it must be relativised to either or both of (i) an evolutionary etiology, (ii) intentional design. But characterising norms in this way requires giving up normativity as a fully causally (better here: dynamically) grounded property, since it involves relativisation to a non-causal (or at least non-causally characterised) condition. This is clear in the case of relativisation to intentional design, because appeal must be made to the designer’s purposes in constructing the system to distinguish what is genuine control. And while an evolutionary etiology is itself causal, and this fact enables holders of the relativisation thesis to maintain a naturalist stance, relativisation to an evolutionary etiology does not provide a causal distinction between merely causal processes and normative processes; rather, to distinguish between the two we must look into the evolutionary past of the systems in question and determine what the systems were selected for. But the past can influence the future only via its effects on current state, and such current state considerations are blocked here. In neither case are we able to appeal to the actual causal structure of the system. The etiological approach, then, ends up destroying the very result that it aims toward: a naturalistic account of function. In the present interactive-constructive context this non-causal approach to naturalising norms is both objectionable and unnecessary. Rather, the appropriate relativisation determining normativity is to the conditions of autonomy of the system in question, and this in turn is a fully dynamically specified property of the system.

Fundamentally normative processes involve two key conditions: (i) a capacity to perform directed work which produces organised dynamical behaviour and (ii) a normative ‘perspective’ which differentiates across outcomes, such that there is a
II: Autonomy

genuine level of system organisation which discriminates between differing types of behaviour of the system as a whole (paralleling the situation in classical control whereby for S+C there is a specifiable goal condition R). Autonomous systems realise condition (i) in the form of the directive structure which generates system-distinctive processes, and condition (ii) in the form of the overall cohesion of the system. Both of these features are genuinely causal, dynamically grounded, properties. They are intimately related in the dynamics of autonomous systems; the directive constraints of the system generates the process patterning that produces the system’s global cohesion, and the degree of system cohesion measures the success of the directive organisation. In sufficiently complex autonomous systems various aspects of the degree of cohesion can become the basis of explicit signals within the system itself, e.g. pain signals, and this internal organisation is the root of the human experience of normativeness.

It is plausible to speculate, though it is difficult to prove, that no non-autonomous system meets both conditions for normativity, except derivatively. Classical control systems, e.g., are explicitly normative (see above) but are not autonomous (though autonomous systems may incorporate control processes); instead they are functionally specified relative to the purposes of the designer/user. Thus, the normativity of a classical control system is derived from that of its (autonomous) creators. Autonomous systems, in contrast, generate their own normativity.

One might worry that there is a circularity in this account inasmuch as the characterisation of autonomous systems is partly in terms of a normative perspective, whilst autonomy plays a central role in the characterisation of normativity. However the close connection is grounded in the dynamical nature of the systems in question rather than being the result of a definitional circularity, and on fine grained analysis the circularity evaporates. Proto-versions of conditions (i) and (ii) can be found which occur independently of each other and in the absence of autonomy. In the case of (i), the proto-condition occurs where there are directive constraints which create organised dynamical behaviour without generating and forming part of a system
II: Autonomy

which is globally cohesive. A rolling boil is a case in point: the boiling structure is cohesive, e.g. it is reconstituted after perturbation by stirring, but the heat source and sink which create the directive temperature gradient which produces the rolling boil are cohesively independent of the behaviour of the boil itself. If the rolling boil disappears because of an influx of cold water there will be no effect on the heat source, which is why the thermal currents will become reestablished and the distinctive cell structure of the rolling boil will re-form. The rolling boil is both cohesively and directly a dependent system. Moreover, although the heat source determines the behaviour of the rolling boil, there is no sense in which the heat source has any reference goal $R$ for the rolling boil’s behaviour in either the classical control or autonomy sense because there is no indexed feedback modulating the behaviour of the boil, and the heat source plus boil do not form a natural supersystem which constitutes a globally cohesive system. In other words there is no level of global organisation which includes both the heat source and rolling boil with respect to which both are constrained.

The proto-version of condition (ii) occurs in systems which are globally cohesive but possess no self-generating directed processes. Many systems meet this condition, including rocks. Like the rolling boil, the cohesion of the generative processes which produce the particular organisation of the rock are independent of the rock itself, but unlike the rolling boil a rock has a high level of endogenous cohesion and can persist long after the generative processes which have given it its organisational structure have ceased. The cohesion of the rock does generate a kind of global dynamical systemic perspective in the sense that the system behaves as an integrated whole in its interactions with the environment. However, as with the rolling boil, there is no real normative structure to the system’s organisation because: (a) as pointed out above, the cohesion of the rock is determined through local interactions rather than through global organisation, and (b) the generative processes (e.g. magma cooling and weathering) and the organised system (the rock) do not form a cohesive supersystem, and precisely because of this there is no unified level of system organisation with respect to which the generative processes are constrained.
Now contrast each of these cases with an autonomous system such as a living cell. Like the rolling boil and unlike the rock, the organisation of the cell must be constantly maintained, but unlike both the rolling boil and the rock the generative processes which produce the cell’s organisation are internal to it, in the form of DNA-RNA-protein synthetic pathways. Thus, the cell meets condition (i). In addition, the cell is globally cohesive and so meets condition (ii). Conditions (i) and (ii) are not merely co-present in the cell, they are mutually supportive: as pointed out above, if you slice a cell in two and thereby disrupt its cohesion conditions it will almost certainly not reform (unlike the rolling boil), and this is because in the case of the cell there is a unified level of system organisation with respect to which the generative processes are constrained. Condition (ii) in effect insists that the dynamical behaviour of the system be evaluated with respect to a set of global conditions, and in the cell this occurs because the existence and functionality of the cell’s DNA and biosynthetic pathways is maintained by the cohesion of the organised cell which results. If the functioning of these generative processes is disrupted sufficiently that a viable cell is not produced, the processes will quickly dissipate. Thus there is a genuine level of mutual constraint between the cohesion of the system and the directed processes which generate that organised cohesion, and consequently there are specifiable normative conditions for the behaviour of the system. In any particular case these conditions will have a biochemically (very complex but) precise content; in general the normative condition is that the directed processes of the system must produce a globally cohesive system which in turn generates the system’s directive organisation. So, another way of describing autonomy is as the co-presence and mutual support of dynamical conditions (i) and (ii). Thus, normativity and autonomy emerge simultaneously, and are in fact aspects of the one dynamical phenomenon.

1.4 Modulatory processes and hierarchical organisation

Autonomous systems may improve the context-sensitivity of their processes and obtain enhanced resilience against potentially disruptive perturbations through
II: Autonomy

modulatory processes such as compensatory feedback. For living organisms, e.g., the impact of declining nutrient intake may be compensated by any one or more of a number of methods for decreasing demand by decreasing activity levels (including hibernation) and increasing supply: switching food sources, increasing nutrient extraction, increased searching and moving to new environments. By contrast, neither such sub-autonomous systems as the candle flame, nor such simple autonomous systems as the catalytic bi-cycle system, possess any modulatory processes. Basically normative modulation tunes key system parameters so that they satisfy system constraint conditions. Thus, in the case of compensatory modulation corrective feedback is applied to the fundamental process cycle of the system if the system drifts outside the range of parameter values necessary for its persistence. Even very simple autonomous systems may possess elementary modulatory processes, and complex multi-parametric modulation is the hallmark of strongly autonomous systems such as organisms.

Modulatory processes introduce high order directive constraints to system behaviour. This provides a partial hierarchical structure to system organisation. As is argued in subsequent chapters, high order modulation is crucial to adaptability, intentionality and learning, since high order constraints open up the capacity to exploit low order plasticity for adaptive self-modification. However, it is as well to note that this hierarchical organisation will be possessed of considerable complexity, since for many adaptive processes there will not be a clear distinction between ‘modulatory’ activities and ‘normal’ activities (strong hierarchical distinctions must be based on relatively clear process interaction asymmetries) and modulatory processes can vary continuously along a number of normatively significant dimensions, including: (1) whether the activity modulated is of peripheral or core significance to the system’s autonomy (i.e., the global entrenchment of the process), (2) the degree of localised interconnectedness of the modulated activity (localised entrenchment), and (3) how sensitive the modulation itself is in respecting the wider functional relationships of the system.
II: Autonomy

With respect to (1), some processes perform activities which, though they assist in generating the system’s autonomy, are themselves derivative or ‘piggyback’ functions. They may improve the efficiency, adaptability, robustness etc., of the system’s fundamental process cycle, but because of their relative superficiality they may be lost with non-fatal reductions in the system’s overall capacity to remain autonomous. For example, colour vision is a functional system which may be lost without resulting in death (at least in many cases). Such modulatory processes may be very highly articulated but, in relation to the global functioning of the system, of fairly superficial importance. In contrast, the modulation of fundamental system processes might be less articulated but nonetheless be of higher global functional significance. For example, failure of the sympathetic and parasympathetic autonomic systems is usually fatal. With respect to (2), a process may be entrenched in the sense that many further system-processes directly depend on it, and hence changes to it will be constrained with respect to the effects it has for these further processes, whilst the complex of processes it is related to may as a whole be functionally superficial. This is true, e.g., of cone-ganglia retinal processes in colour vision. The converse may be true – a process may be directly only connected with a few other processes but nonetheless perform a fundamental system function. Dimension (3) involves two subdimensions: (3a) how widely dependent the modulation is upon the states of other processes (synchronic conditionalisation), and (3b) whether the parameters of modulation have themselves been determined as a function of learning (diachronic conditionalisation). Dimension (3) concerns how informed the modulation is, and this can be achieved by direct causal connectivity to other system states (e.g. when a trauma induces a general conservatism in an animal’s behaviour pattern), or by apt structuring of the modulation based on a learning history (e.g. a conditioned pain reaction). In the latter case the modulation may implicitly respect functional relationships, without being immediately conditionalised upon them, because of historical structuring.

Each of these factors will play a role in determining the hierarchical organisation of the system, which will consequently be very complex. Ultimately the hierarchical
II: Autonomy

organisation of a system’s process relations is determined by the relationships between process activities and the cohesion of the system as a whole, and there can be a fairly continuous variation in the domain and range (degree of specificity/generality) of the modulation, such as the way feedback may range from momentary correction to tactical revision to strategic revision (to starting all over again in the next life).

1.5 Interaction and the locus of directive constraint

Autonomous systems are dynamically coupled with their environment, and the organisation of their constitutive processes results from a complex and subtle interplay of internal and external factors. For any given process one may identify the locus of directive constraint of the various factors which contribute to the process according to whether they stem from the system or environment. When gardening Max is controlling (or attempting to control) part of his environment (internal locus of directive constraint), when listening to speech he is relinquishing at least some control of internal state to the speaker part of his environment (locus external), when frolicking in the surf he may be simply interacting with his environment (no locus of directedness either way), but when making sexual love with his long-term partner he may initially be importantly engaged in a delicate process of considerate mutual control (directive locus partly internal and partly external). All these distinct conditions grade back through the panoply of autonomous systems, and also shade continuously into one another. (Later Max’s love-making condition, e.g., may shade into, and culminate in, unmodulated interaction.) These distinctions are not arbitrary, they are based on the presence or not of real modulatory processes. A plant growing phototropically, or a human responding hormonally to day length, is simply in open-loop interaction with sunlight in its environment because in each case there is simply a conditional response to sunlight intensities, but the human, unlike the plant, has the

---

8There is here no representation of a reference condition nor construction of, and response to, deviation signals therefrom. The tips of phototropic plants produce a growth factor, known as auxin, at constant rate. Auxin is transported down the vascular system to the shoot, where it stimulates cell growth. Photosensitive molecules in the tip regulate transport of auxins into the plant’s vascular system, probably by altering membrane permeability to auxin. In conditions of no light auxin transport is laterally evenly distributed, and the plant grows vertically.
II: Autonomy

capacity to actually control for light exposure.

The plant phototropic response adaptively solves the problem of efficient growth without having to explicitly represent it or construct a solution. In these cases there is simply an open loop interaction between the system and the environment: if the stimulus is present in the environment the system reacts and if not, not. However the greater human capacity for internal directedness confers greater adaptability and this too can be advantageous despite the additional organisational complexity required to support it. Here the system modulates its processes using internally generated signals (e.g. hunger), ‘short-circuiting’ the open loop to switch it on and off: The system is still sensitive to signals from its environment, but the relation between environmental signals and behaviour is mediated by layers of internal directive constraint which can (at least potentially) ensure that behaviour is more aptly related to the system’s own requirements. Notoriously, chickens will continue to eat until they kill themselves. Here the eating activity is a simple open-loop behaviour. This strategy is satisfactory so long as food does not often occur in concentrations sufficiently high to overload the chicken’s ability to digest it; however in environments where this is not the case the strategy becomes dysfunctional and it would be useful for chickens to have a satiation threshold which suppresses the behaviour. This would add a conditionalisation to the behaviour – eat whenever there is food, unless the satiation threshold has been exceeded. The conditionalisation represents a modulation of the signal-behaviour sequence by an internally generated signal which is correlated with an autonomy-relevant functional state (dangerously high food ingestion). This modulation tailors the activity more precisely to the system’s requirements.

Variations along the dimension of locus of directedness have significant implications for the wider organisational properties of the system. Some degree of internal directedness is necessary for autonomy, and even comparatively simple autonomous

---

However light striking the tip of the plant causes the production of molecules which inhibit auxin transport, with the result that auxin concentration becomes higher on the side of the tip away from the light, where transport is not inhibited. Because of this more auxin is transported to growing cells in the side of the shoot furthest from the light, causing this side of the shoot to grow more quickly than the lighted side. The end result is that the shoot tends to bend in the direction of the light source.
II: Autonomy

systems such as cells have significant degrees of internal control (e.g. the cell’s cycle clock which synthesises environmental signals to determine the cell’s phasic behaviour). However autonomous systems must also regulate their processes context-sensitively in order to be adaptive. Unmodulated but apt external sensitivities can be an efficient and powerful form of adaptive behaviour if they implicitly embody significant facets of the system’s constraint relationships (e.g. the case of plant phototropism cited above). Typically, even the most sophisticated organisms will possess some mix of open-loop and internal directed processes, and will be subject to selective forms of shaping from their environment. In consequence, autonomy is a multi-dimensional, graded, and context-dependent property. Autonomous systems will have a complex autonomy profile with dimensions corresponding to the hierarchical, internal and external directive constraints involved in the system’s organisation. Moreover this autonomy profile may shift (at various rates along various dimensions) over the lifespan of a system: as a system matures its autonomy can increase along some dimensions (e.g. functional adaptedness, cognitive sophistication) and decrease along others (e.g. physiological resilience, developmental plasticity).

Further complexity arises in distinguishing functional internal subsystems from pathological states, such as cancers, and competing systems, such as parasites. This issue is complex and shall not be explored here other than to note that, once again, cohesion is the appropriate yardstick for making the appropriate distinctions. Normal cells constitute subsystems of autonomous multicellular organisms because the conditions of their cohesion reinforce the cohesion of the organism (and vice versa). Conversely, cancer cells are pathological because their conditions of cohesion undermine the cohesion of the organism (though not, unfortunately, vice versa; with respect to its local context, a cancer cell is just an autonomous system like any other). Likewise for parasites, where we find examples all the way from the strongly autonomous, like lice, to the tenuously autonomous, such as those viral parasites that must even rely on their host cells for reproduction.
II: Autonomy

2. Anticipation

2.1 Anticipative Closure: General Specifications

The root sense of anticipation is of a feedforward process with a telic closure condition (cf Rosen 1985a). This model of autonomy allows us to define anticipative closure: A system achieves anticipative closure when the interactive consequences of an action A in an environment E have significant positive autonomy value for the system. The hunger $\rightarrow$ hunt $\rightarrow$ eat $\rightarrow$ satiation sequence achieves anticipative closure for an adapted system in its niche environment, the effect of the anticipatory hunting + eating action is to close off the initiating signal while at the same time resetting the organism’s internal condition to a more viable one. More formally, for a particular environment E an autonomous system’s anticipative closure conditions are given by the specification of those system organisational features which support system functional capacities that in E give rise to the actions A that have sufficiently positive autonomy value. Over prolonged time scales, anticipative closure is achieved when the sequence of actions performed during earlier time-periods cause the system to remain viable over later time-periods. In the simplest cases the anticipative closure conditions are entirely implicit in the system’s operation, the system itself having no explicit internal access to them. In more complicated systems employing degrees of internal control some aspects of the closure conditions can become explicit.

This account of anticipation is still too abstract to tell us very much about the nature of a system’s anticipatory closure conditions, a particular problem being that it is temporally and operationally unbounded. When a lion hunts is it anticipating eating to remove its present hunger signal only, or this and the next that would normally have arisen, or all future hunger? Is it also anticipating acquiring the energy to mate? To mate and successfully defend its mate, or a large pride, or ...? And so on. Without a natural way of specifying bounded time-periods for anticipative closure the account threatens to attribute to all autonomous systems the hope for eternal life, with the result that system actions which achieve anything short of this constitute failure.
II: Autonomy

Similarly, without a way of specifying the operational context of autonomous system actions, the account threatens to attribute the hope for omnipotence to even the most mundane activities. Effectively, without some way of specifying natural boundedness the account of anticipative closure remains vacuously inclusive. However we can get some handle on the issue of natural boundedness by examining in more detail the processes involved in autonomy.

Actual autonomous systems achieve viability by means of physical processes which frequently have a cyclic structure and which must be completed for the system to persist. For example, a cell has numerous metabolic cycles (e.g. the Krebb cycle) which allow it to complete its major meiotic or mitotic cycle; an organism has a hunting, feeding and resting cycle of activities, which enable it to complete its internal reproductive cycles (and contribute to completing its sexual reproductive cycle, if it has one that requires its action). And so on. Particular actions by the system will play some role (often multiple roles) within these cycles, and the anticipative closure conditions for the actions are determined relative to the cycle or cycles of which they form part. Of course, not all processes in autonomous systems are periodic and thus possess concrete closure conditions. Some adaptive processes, notably those involved in development and learning, may have aperiodic, chaotic, or indefinitely transient dynamical trajectories. In these cases the content of the anticipation may be inherently vague, being determined by the nature of the relations between the dynamical evolution of the processes and the wider functioning of the system. When a learning system expends energy to learn something, e.g., it need not be for some immediate specific purpose, but either for future general use – as when lion cubs practice fighting – or just in case ‘it might come in handy’, as when a mouse might explore and remember a hole. In general, the specificity of the anticipations is determined by the specificity of the process. For example the closure conditions for mating are quite specific (reproduction) whilst the closure conditions for general learning are vague (acquire an adequate group of discriminatory/behavioural responses, whatever that might come to, for the domain in question). Overall then, the temporal and operational boundedness of an autonomous system’s
II: Autonomy

 anticipations are determined by the characteristic cycle-periods and contexts of its constitutive processes.

2.2 Anticipative Closure: Detailed Specifications

Identifying the natural boundedness of anticipative closure conditions is a step towards presenting a physically interpretable theory of anticipation, but it is necessary to outline more precisely the factors involved in successful/unsuccessful feedforward action. Whether the actions of an autonomous system cause it to remain within its viability envelope is contingent on two broad sets of factors: (i) It depends on the actions of the system being functionally coherent (i.e., properly articulated and sequenced). For example, it is no good if, in attempting to perform an action — say, picking up a pen — you innervate a particular muscle group and simultaneously innervate an antagonistic muscle group — the antagonistic group must be inhibited if you are to be able to complete the action. (ii) It depends on the environment possessing the right properties — the interaction must produce the appropriate input to the system’s processes (e.g. nutritious, purchasable food). Furthermore, each of these factors has both a global and a local context: the system must be functionally coherent both in terms of its overall process structure (e.g. the held pen must be usable to support autonomy, say by writing a cheque) and in terms of its detailed interactions (e.g. the system must be able to write in the circumstances to hand), and the environment must be of the general kind capable of supporting the system (e.g. nutritious purchasable products), and also possess the specific properties which will allow the system’s immediate actions to be successful (e.g. a cheque-using payment system). Each of these conditions could fail to obtain: the system may be globally dysfunctional and/or unable to properly perform or sequence its activities, the environment may not be such that the actions the system performs result in the input required by the system’s processes.

To restate these points more systematically, there are two major dimensions to the anticipative closure conditions of autonomous systems, and each of these dimensions
II: Autonomy

has two aspects:

_Dimension 1 (dim1) is that the system’s processes are locally coherent; i.e., they result in a procedurally coherent interaction with the environment. This involves: (i) the sub-process concerned having the appropriate properties to engage in efficacious interaction, and (ii) the environment being of the right type for the process to reach local closure._

_Dimension 2 (dim2) is that the system’s processes are globally coherent; i.e., they result in the system remaining viable. This involves: (i) the system being organisationally coherent overall, and (ii) the environment having the right resources to satisfy the constraints associated with the organisation of the system._

This now provides the concept of anticipative closure conditions enough structure to make preliminary analyses of actual systems. §1.2 briefly discussed hunger as an example of a control problem which humans faced, and contrasted the explicit control process involved in suppression of the hunger signal with the virtual control of nutrition intake. This distinction can now be made more precisely in terms of the closure conditions for hunger-associated behaviour. The anticipative closure conditions for the activities associated with hunger divide into two distinct groups corresponding to the two major dimensions identified above:

**Dim1.** (i) The organism’s hunting & feeding processes will result in the ingestion of food. (ii) The ingestion of food will result in the inhibition of the hunger signal.

**Dim2.** (i) The organism is capable of converting food into usable products for its metabolic processes. (ii) The food obtained is in fact nourishing, and hence eating will in fact prevent starvation.

The first group involve the strategies used for obtaining food and the mechanics of eating the food. The overall closure condition for this group is that the system will
II: Autonomy

successfully suppress its hunger sensation. The second group of closure conditions relate to the nutritional value of the food obtained, and the overall closure condition for this group is that in satisfying its hunger the system will be adequately nourishing itself. That is to say, the second dimension of the closure conditions for eating is that eating should prevent starvation.

3. Modelling autonomy dynamically

Thus far autonomy has been characterised qualitatively, and as such the account is a partial guide rather than a rigorous model of the phenomenon. Because of the sheer complexity of autonomy, attempts at rigorous modelling may prove impractical (and potentially counterproductive if important dimensions are suppressed to achieve tractability). However, ultimately autonomy must grounded in a dynamical theory, and to conclude this section some preliminary considerations are introduced for developing such a theory.

As a framework for what follows, it is supposed that systems S are specified dynamically by some S-tuple of dynamical parameters p(S), the general relations among the p(S) specifying the category of system and an S-tuple of the p(S) values the specific system state. Similarly, it is supposed that environments are specified dynamically by some E-tuple of dynamical parameters q(E), the general relations among the q(E) specifying the kind of environment and an E-tuple of the q(E) values the specific environment (the specific environmental state). Talk of “environment E” means that kind of environment, including all of its specific quantitative variants; talk of many environments means many kinds of environments. (In common parlance there is one environment for a system: everything else there is, the total environment; nevertheless, it is also convenient to speak of the many relatively distinctive and stable conditions within this environment as themselves kinds of environments and this practice is used here.)

In characterising autonomous systems it is necessary to pay attention to the way in
which the system’s interactions satisfy its cohesion requirements. The relationship between these two sets of factors determines what will be termed the system’s constraint envelope. The constraint envelope for \( S \) in \( E \) is comprised of the system viability envelope \( V(S) \), the set of internal conditions for which \( S \) is viable, and the system-environment interaction envelope \( I(S, E) \), the set of correlated system and environment conditions that result from interaction between system and environment. Here the focus is on the viability envelope. The full constraint envelope will play an important role in understanding adaptiveness (chapter III).

The basic intuition behind the concept of a viability envelope is that of the interrelationships amongst the constitutive parameters of an autonomous system.

\[ V(S) = \text{the set of all S-tuples, each S-tuple comprising a complete set of joint values of the system dynamical parameters } p(S) \text{ for which the system is viable.} \]

Consider a system \( S \) described by the change in time of a set of parameters \( a, b \) and \( c \). A typical type of dynamical analysis of such a system would be to hold \( a \) & \( b \) fixed, and look at the way changes in \( c \) affect system behaviour. But this misses what is distinctive about the interrelations in autonomous systems. Typically \( a, b \) and \( c \) will be organisationally interdependent; for example, outside a certain range of values for \( c \), \( a \) and \( b \) may no longer hold for the system (the system is no longer a system, or at least the system it was). \( a, b \) & \( c \) may each hold for the system only under certain conditions, which include the values for the remaining two parameters falling within some range. In these conditions there arise mutual constraints amongst the parameters imposed on the coherency, not just the quantitative behaviour, of the system as whole.

Call \( C_a \) the statement that parameter \( a \) holds for \( S \) and that \( b \) is within its required range, \( b_a \) say, for this to happen, and similarly for \( c \). There will be corresponding statements \( C_b \) and \( C_c \) for \( b \) and \( c \). These ranges need not coincide \((b_a \neq b_c \text{ etc.})\); however, since in general strong levels of parametric interdependency can be expected (see below), for simplicity only the system states where they do coincide will be considered, so that there is a common set of ranges, \( a_v, b_v, c_v \) say, which
II: Autonomy

appear in the C statements. Join to the C statements all statements of the form CR\(_{a,b,c}\) which state quantitatively conditionalised qualitative interdependent variation relations among \(a\), \(b\) & \(c\) in the conditions where the C statements hold, e.g. that if \(c = n\), \(n \in c_v\), then if \(a\) increases then \(b\) decreases. The CR statements describe the panoply of feedforward and feedback relations that obtain in the system; together they specify, in a state-dependent way, the qualitative way the system will respond to changes of state. This captures, e.g., the relations determining the stability, and hence resiliency, of the system autonomy at each state where \(a_v\), \(b_v\), and \(c_v\) hold. Finally add CPH, the statement of any further particular physical and historical constraints associated with the instantiation of the system. Call the resulting conjunction of C, CR and CPH statements \(C_s\), the joint coherency conditions characterising the system.

The coherency conditions \(C_s\) constitute the identity conditions of the system, the conditions under which it will retain its integrity. \(C_s\) is satisfied if and only if \(S \in <a_v,b_v,c_v>\). Finally, the condition \(S \in <a_v,b_v,c_v>\) picks out a set of dynamical states in S’s state space (not necessarily continuously connected) which are thus associated with a particular \(C_s\); call this set \(DC_s\).\(^9\) Cases of particular interest occur when, e.g., \(DC_s\) contains a region of modest resiliency connected by a narrow ‘bridge’ with a region of higher resiliency and the system’s best adaptive strategy is to successively modify itself so as to move from one to the other under its coherency constraints – this is typical of all cognitively driven adaptability and many other cases as well.\(^10\)

---

\(^9\)The bundle of constraint relations \(DC_s\) is not isomorphic with state space since whatever relations there may be among the members of the bundle, the relations will order them in terms of systemic coherency rather than dynamical contiguity. Several state space points may not differ in their constraint characterisation and so be associated with the same \(DC_s\). Conversely, nearby state space points may differ radically in their constraint characterisation. In this case a small distance in dynamical state space corresponds to a large difference to constraint relations (\(DC_s\) profile).

\(^10\)Though it is not possible to specify a distance comparison between two constraint bundle members, as it would be for two state space points, it is possible to compare the structural relations between the two constraint profiles and, together with the state space portrait, specify a physical transformation which will move the system from one \(DC_s\) profile to the other. Note that interestingly complex constraint structures do not emerge until we come to model systems which are autonomous. Systems which are arbitrarily defined according to observer interest and systems which possess natural identity conditions but which are merely cohesive don’t have interesting interrelations amongst the parameters which constitute their identity conditions. As was argued in the text above, in the case of merely cohesive systems the identity conditions of the system do not impose constraints on its cohesion conditions (though the converse is not true – the cohesion conditions of the system do impose constraints on its identity conditions). Consequently the constraint structures of such systems are null or trivial. It is one of the defining characteristics of autonomous systems that their identity conditions impose strong constraints on their dynamical/physical character.
II: Autonomy

DC\(S\) denotes a subregion of state space \(<a,v,c,v>\) in which there is a phase separation—an emergent locally coherent system of interacting elements differentiated from the environment. \(a, b \& c\) will be the constitutive parameters governing the phase separation, and \(C_S\) describes the cohesion conditions of the region \(<a,v,b,v,c,v>\). If the system is resilient then it will be stable against perturbations and tend to remain within \(<a,v,b,v,c,v>\). For this to be true the dynamics must be such that certain parameter values are mutually reinforcing, inasmuch as particular parameter value combinations tend to stabilise the system away from the boundaries. Assume that within \(<a,v,b,v,c,v>\) there is a ‘sweet spot’—a region in which the system is maximally resilient against perturbation. Assume further that this resilience declines away from the sweet spot. The resilience pattern may be complex, but in general resiliency will tend to be low near the boundary. At the boundary itself it will be zero. \(DC_S\) will include a description of the features of this resiliency pattern and specify the relationship between the sweet spot and the cohesion conditions of the system.

The resiliency pattern of an autonomous system will constitute a dynamical attractor basin in its state space. With respect to the state space portrait of a given autonomous system, some of the points of the space will correspond to system states in which the coherency amongst the system’s processes is low and these will be states which possess poor resilience (e.g. the poor social skills of a maladjusted chimpanzee which result in increased stress and malnourishment from diminished access to food resources). Term these conditions *constraint mismatch*, because the constraints which pertain to the functioning of the system’s processes are in conflict. Sometimes conditions correspond to system states where there is relatively high coherency amongst system processes (e.g. a healthy lion’s successful hunting strategy is providing ample food to fuel the lion’s physiological processes); these are *constraint match* conditions.

However the state space description *isn’t* the most perspicuous level at which to analyse the constraint structure of the system—it is necessary to know *why* the region \(<a,v,b,v,c,v>\) has the structure it has, why it is resilient, when it is, and so on. Why, for
example, is a critical aspect of the behaviour of S? It is critical because it is a member of \( C_b \) & \( C_c \) (just one would be sufficient). This kind of question abstracts away from the details of specific state space points and looks instead at regional relations. In other words, the appropriate level for understanding constraints is in terms of the structural relations among \( C_a, C_b \) & \( C_c \) which jointly make up \( C_S \). These relations are implicit in the state space portrait of the system, and are made explicit in the constraint relations portrait. (Conversely, the detailed dynamical information which is explicit in the state space portrait of a system is partial and implicit in its constraint relations portrait.) This kind of constraint information allows us to predict negative and positive feedback cycles in the system’s dynamics. For instance, if process \( A \) can only function under the set of conditions \( C_A \), and process \( B \) only functions under the set of conditions \( C_B \), and the functioning of \( B \) contributes to \( C_A \), then \( A \) is partially dependent on \( B \). If \( B \) comes to function in a fashion which does not contribute to \( C_A \) then \( A \) may cease to function. If it is also the case that \( A \) contributes to \( C_B \), then if \( A \) ceases to function \( B \) may also cease to function. This debilitating positive feedback cycle results from the initial mismatch between \( B \) and \( C_A \) which feeds through to cause a mismatch between \( A \) and \( C_B \).

These constraint relations are determined by the process organisation of the system in question, plus its physical constraints – in fact \( DC_S \) will be the intersection of these two sets of constraints. It is necessary to be careful here: a system’s organisational structure per se is not the same as – or fully determinative of – its constraint structure. To properly specify \( DC_S \) requires both an organisational and a physical characterisation of the system. Constraint representations are neither wholly functional, in the usual sense, nor wholly physical forms of explanation. Rather, they straddle the divide between the two, systematically relating both sets of factors. There are all sorts of ‘physical’ characteristics of cells, such as the difference in size between the hydration spheres of sodium and potassium ions, which generate systemic constraints (e.g. on the types of membrane spanning proteins which can function as sodium-specific or potassium-specific channels, on the conductance of such channels, and on potential membrane capacitance) but which are not in
II: Autonomy

themselves specifically organisational. A specifically organisational constraint, by contrast, would be something like: ‘In a four stage linear or monoplex catalytic process the substrate for the third stage must be a product of the second stage, and some product of the third stage must be a substrate for the fourth stage’. Physical constraints certainly have organisational implications (e.g. via the functional implications of the actual capacitance of the membrane), but when trying to analyse the overall set of systemic constraints it is necessary to remain aware that this characterisation cannot be in exclusively organisational terms – specific organisational features may reflect physical constraints.

4. Conclusion

This chapter has developed a theory of autonomy designed to model the essential organisational features of living systems. Autonomous systems have been characterised as self-generating cohesive systems, and some of the principle organisational dimensions of autonomy – which play a role in generating the rich variety of autonomous systems – have been discussed. Autonomous systems can possess various forms of modulatory processes and quasi-hierarchical organisation, and these play a critical role in determining the locus of directedness of the system’s interaction processes, which can vary from externally directed to self-directed. The relationships between the various forms of autonomous organisation and autonomy are further explored in chapter III, and self-directedness is used in chapter IV as the basis for a constructivist account of intelligence and learning. The analysis of process closure in autonomous systems represents the root conceptual basis for understanding normativity in the I-C framework, and it plays a fundamental role in all further work in this thesis. In particular, it will be used to ground accounts of anticipation, intentionality, and the adaptive constraints on self-directed learning developed in chapters IV, V and VI. Some preliminary considerations for modelling autonomy dynamically have also been discussed, and this analysis further articulates the structure of autonomy and suggests ways in which the concept might be developed.