Chapter 3

We are the intelligence that preceded us in its new material representation - or rather, we are the re-emergence of that intelligence, the latest embodiment of its struggle for survival.

Fred Hoyle, *The Intelligent Universe*

Emergence

3.1. Overview

In this chapter the notion of emergence is investigated. The aim is to clarify the meaning of the concept and examine its formulation in the context of a computationalist metaphysics in preparation for the development of the unifying framework of emergent artificiality described in chapter 5. First, a popular definition of emergence is presented and the historical background to the idea is given. A problem associated with the notion of self-organization having bearing on the issue of emergence is identified. Basic system-theoretical concepts associated with the modern formulation of emergence are introduced and phenomenological issues (concepts of emergence and types of emergent) are discussed. A formal framework for emergence, the distinction between emergent computation and computational emergence, and the nature of emergence in cellular automata are briefly examined. Finally, two philosophical perspectives associated with the concept of emergence are identified involving consideration of (i) epistemological issues (intrinsic and extrinsic observation, observational relativism, reductionism) and (ii) ontological issues (open and closed emergence, downwards causation, pluralism).

3.2. What is *Emergence* ?

In this section, the motivation underlying a consideration of the notion of emergence is briefly described and the concept of emergence is defined.

3.2.1. Motivation: The Unification of Artificiality

It may be argued that it is the inability to deal with what are considered the `hard' or defining problems of artificial intelligence (AI), for example, situated-action, computer vision etc, that has led to the displacement of top-down symbolic (or explicitly-representationalist) approaches by bottom-up or connectionist strategies. According to Langton (1989b), however, this trend or `paradigm shift' (chapter 1) does not extend across *artificiality* (chapter 4), that is, the sciences of the artificial (Simon,81); the discipline of artificial life (A-Life) adopts a bottom-up approach *by definition*. Furthermore, and following this definitional precedent, the bottom-up approach has been applied to other artificial sciences: For example, artificial culture (Gessler,94) at a level above AI and artificial physics (Fredkin,90) at a level below A-Life, the logical conclusion to this project being artificial reality (A/VR) (Benedikt,91). Applicability of the connectionist approach across phenomenal domains lends support to CA-computationalism (chapters 2 and 5); however, what remains unclear is how such phenomena are to be unified given the assumption that the world is a coherent whole (chapter 1). Various schemes for unifying such disciplines are conceivable and are examined in chapter 5 where a unifying framework for artificiality (that is, artificial analogues of natural phenomena such as matter, life and mind) is presented. In preparation for the discussion in that chapter, it is necessary to investigate a concept that is central to the bottom-up approach, and which provides a means for unifying artificial domains based on an isomorphism with the phenomenal hierarchy that is assumed to have evolved in the natural world (chapter 1), viz. *emergence*.

3.2.2. Definition: Emergence

Popular scientific and philosophical discourse holds that the essence of the concept is traceable to an observation made by the ancient Greeks, viz. that

"the whole is more than the sum of the parts."

However, this statement is at best incomplete since it does not clarify the distinction (if any) between emergence and related concepts such as holism and self-organization. More importantly, it does not describe the way(s) in which a whole is more than the sum of its parts; that is, it fails to clarify the ontological[1] and epistemological[2] issues associated with the concept. An attempt at articulating a more precise definition of emergence has been made by Mayr (1982) who maintains that

> *emergence* denotes the appearance of new characteristics in wholes that cannot be deduced from the

[1] Ontology refers to the study of the being or nature of things.

[2] Epistemology refers to the study of how things are known.

> most complete knowledge of the parts, taken separately or in other partial combinations. (p.63)

This definition is clearly an improvement on the original maxim since it addresses both the ontological and the epistemological issues associated with the concept. Appreciation of this dual perspective is crucial since it provides a means by which the concept of emergence can be examined within the scheme of ontic (productive, organizational) and epistemic (interpretative, observational) relations between *systems* (section 3.4.1) and the *anthropic component* (human artificer-interpreter) briefly introduced in chapter 1 and described more fully in chapter 7. However, the above statement clearly does not constitute a formal definition of emergence. In an attempt at constructing such a definition, various philosophical schemes which provided the basis for modern formulations of emergence will be briefly discussed in order to trace the development of the concept to its modern systems-theoretical forms. The objective is to examine the various issues associated with the concept and thereby provide a basis for the critique of the unifying framework of computationally emergent artificiality presented in chapter 7.

3.2.3. Historical Background

It is to Chapter VI of J.S.Mill's *A System of Logic Ratiocinative and Inductive* (1843) that perhaps the first implicit reference to the concept of emergence in a modern scientific context can be traced. Mill describes what he calls *heteropathic laws* as "laws of combined agency [which] are not compounded of the laws of the separate agencies" (p.375) and goes on to state the following:

> as a general rule, causes in combination produce exactly the same effects as when acting singly: [however,] this rule, though general, is not universal: that in some instances, *at some particular points in the transition from separate to united action, the laws change, and an entirely new set of effects are either added to, or take the place of, those which arise from the separate agency of the same causes*[3]: the laws of these new effects being again susceptible of composition, to an indefinite extent, like the laws which they superseded [emphasis added]. (p.376)

However, the earliest documented use of the term `emergence' is attributed to the philosopher of science George Henry Lewes in the middle of the nineteenth century (Kenyon,41). Lewes anticipated the epistemological conclusions drawn from modern research into the behaviour of complex *non-linear*[4] dynamical systems by distinguishing

---

[3] This concept of emergence is consistent with both cumulationist (or inclusivist) and non-cumulationist (Bunge,60) emergentism as described in chapter 5.

[4] *Linear* systems obey the *superposition principle* since they are decomposable into independently analyzable components and composition of understanding of the isolated components leads to *full* understanding of the system. The principle does not hold for *non-linear* systems since in this case, primary behaviours of interest are properties of the *interactions* between components as contrasted with properties of the components themselves; isolating the components *necessarily* leads to the disappearance of interaction-based properties

between *resultants* and *emergents*: In the former, the sequence of steps which produce a phenomenon are traceable whilst in the latter they are not. Thus, according to Lewes, emergence is a statement of the *epistemological* limitations of phenomenal observation. In the twentieth century, the concept of `emergence' was formally introduced by Conroy Lloyd Morgan in *Instinct and Experience* (1912) and a theory based on the concept was developed in his *Emergent Evolution* (1923). Like Lewes, Morgan regarded the distinction between resultants and emergents as inductive and empirical, and not metaphysical. Emergence was understood as implying that the higher orders of being produced in the evolutionary process were not mere resultants of what went before and were not contained in them as an effect is in its efficient cause; hence, emergentism was not equivalent or reducible to mechanism. Blitz (1992), who has extensively researched Morgan's philosophy of emergent evolution, maintains that it consists of three premises:

(1) evolution is a universal process of change productive of qualitative novelties

(2) qualitative novelty is the emergence of a property not possessed by any of its parts

(3) reality can be analyzed into levels consisting of systems characterized by emergent properties

However, Morgan's scheme was purely descriptive and did not provide an explanation for the phenomenon of emergence (Collingwood,45). It was in Samuel Alexander's *Space, Time and Deity* (1920) and J.C.Smuts' *Holism and Evolution* (1926) that the first attempts at an explanatory framework for emergent evolution were made. On Smuts' view of emergence, nature is permeated by an impulse towards the creation of wholes; novelty, the defining characteristic in Morgan's scheme, is considered as a merely derivative phenomenon. Alexander, on the other hand, anticipated modern notions of self-organization (section 3.3) and functionalism (chapter 1) by identifying emergence with the tendency for things to arrange themselves into new *patterns* which as organized wholes possess new types of structure and new properties or 'qualities'. The connection to functionalism is implied in the conception that quality depends on pattern (organization, form or structure). Alexander's concept of emergence provides the basis for the unifying framework of computationally emergent artificiality presented in this thesis and a detailed examination of his metaphysics (ontology) will be made in chapter 5 when this framework is described.

3.2.4. Towards A Theory of Emergence

One of the earliest presentations of the concept is that given by Broad (1925) who provides the following outline for a theory of emergence:

- there are certain wholes, composed (say) of constituents *A*, *B*, and *C* in a relation *R* to each other;

(Langton,89b).

- all wholes composed of constituents of the same kind as *A*, *B*, and *C* in relations of the same kind as *R* have certain characteristic properties;

- *A*, *B*, and *C* are capable of occurring in other kinds of complex where the relation is not of the same kind as *R*;

- the characteristic properties of the whole *R(A,B,C)* cannot, even in theory, be deduced from the most complete knowledge of the properties of *A*, *B*, and *C* in isolation or in other wholes which are not of the form *R(A,B C)*.

Pap (1952) clarifies the epistemological character of the above scheme as follows:

> a law correlating a quality *Q* with causal conditions of its occurrence can, without obscurantism, be argued to be *a priori unpredictable* if the predicate designating *Q* is only ostensively definable [that is, via *a posteriori* observation of *Q*]. (p.304)

Hence, "an emergent law is *deducible only a posteriori, or unpredictable a priori*." (p.305) He goes on to state that,

> if to say that quality *Q* (or relation *R*) is *absolutely emergent* is to say that the law correlating *Q* (or *R*) with quantitative physical conditions is *a priori* unpredictable, it follows that absolute emergence is relative to a system of semantic rules. (p.310)

As will be seen in what follows, this position anticipates Cariani's (1989) concept of emergence-relative-to-a-model (section 3.5.1). An alternative scheme is presented by Pepper (1926) who identifies emergence as a kind of change different from that of random changes ("chance occurrences") and mechanistic changes ("shifts"):

> emergence .. is a cumulative change, a change in which certain characteristics supervene upon other characteristics, these [latter] characteristics being adequate to explain the occurrence on their level. (p.241)

Supervenience is the idea that a higher level phenomenon is (causally) dependent on a lower level phenomenon for its existence[5]: For example, mental phenomena are (causally) dependent on the occurrence of physical events in the brain. However, the supervenience relation between a mental event and the associated physical event does not entail a one-one mapping between *substrate* and *emergent* (or supervenient): Many physical brain states can support the same mental state; thus, mind is said to be *multiply-realizable* (chapters 1 and 4). As Klee (1984) states, there are "multiple distinct determinative micro-bases for the *same* macro-property" (p.56) Hence, there is a close relationship between supervenience and functionalism. Pepper's cumulationist formulation implies, however, that emergents (that is, the products of emergence) must be epiphenomenal or non-causal; thus, the possibility of `downward' causation (section

5 Chalmers (1996) defines supervenience as follows: "B-properties *supervene* on A-properties if no two possible situations are identical with respect to their A-properties while differing in their B-properties." (p.33)

3.9.2) is rejected *a priori*.

According to Pepper, the theory of emergence involves three propositions:

> (1) there are levels of existence defined in terms of degrees of integration; (2) there are marks which distinguish these levels from one another over and above the degrees of integration; (3) it is impossible to deduce the marks of a lower level from those of a higher. (p.241)

On this basis, the concept of emergence, originally conceived in terms of the relations between parts and wholes, is readily recast in systems-theoretic terms: For example, Pepper's propositions are directly associated with three issues: (i) complexity (section 3.4.5), (ii) types of emergent (section 3.5.2) and (iii) the limits of reductionism (section 3.8.3). Prior to investigating such issues, however, it is necessary to briefly examine the connection between emergence and a historically antecedent concept, viz. *self-organization*. An appreciation of the notion of self-organization is important because the latter gives rise to a problem which bears directly on the issue of emergence as will be shown in the following section and again in chapter 6.

3.3. Self-Organization: Problems in Emergence

Perhaps the earliest reference to self-organization in the history of Western thought is to be found in the writings of the Ionian Greeks (7th-6th centuries B.C.). For the Greek Anaximander, the undifferentiated substance of reality was `the Boundless' - or *apeiron*[6] - which was spatio-temporally infinite and qualitatively indeterminate (Collingwood,45). The link between Anaximander and self-organization is cosmological, viz. "a world is .. a thing that *makes itself* [that is, self-organizes] wherever a vortex arises in the Boundless [on account of creative processes within the Boundless which take the form of rotary movements]; hence *a world is also a world-maker* or a god [emphasis added]." (p.35) Modern formulations of the concept of self-organization are traceable to the works of the early cyberneticists (Yovits,60,62) (von Foerster,62) (Yates,87), and there is general agreement as to the meaning of this concept. Zwierlein (1984) defines self-organization as "the phenomenon that some (i.e. non-linear dynamic) systems may by irreversible processes spontaneously generate a structure of higher complexity than the original starting point." (p.290) Similarly, Garfinkel (1987) defines self-organization as the general term for the processes by which order and structure emerge and Heylighen (1993) as "a spontaneous (i.e. not steered or directed by an external system) process of organization, i.e. of the development of an organized structure." (p.1) Von der Malsburg (1987) defines self-organization in the following terms:

- There is a *system* (section 3.4.1) consisting of a large number of microscopic elements. The system initially is in a relatively undifferentiated state.

---

6 A Heideggerian interpretation of the concept of the *apeiron* is presented in chapter 6.

- There are self-amplifying fluctuations (i.e. deviations from) the undifferentiated state.
- Some limitation of resources forces competition among fluctuations and selection of the fittest (i.e. the most vigorously growing) at the expense of others.
- Fluctuations cooperate. The presence of a fluctuation can enhance the fitness of some of the others, in spite of the overall competition for resources in the field. (In many systems the `fitness' of a fluctuation is identical with the degree of cooperativity with other fluctuations.)
- Whole systems of cooperatively interacting fluctuations emerge as ordered, differentiated states, or ordered modes - the order often extending over a wide area.

In the context of a discussion of the evolution of biological systems, Erwin (1994) distinguishes between internal, external, and hybrid emergence: In the first, emergence has its origin in processes internal to a system (that is, endosystemic); in the second, emergence is attributed to processes that are external to the system (or exosystemic), that is, in its *environment*; in the third, both approaches are involved in emergence. Identification of which type of emergence (internal, external or hybrid) is occurring in a given system is important since it enables a problem associated with the notion of self-organization to be resolved. This is the *semantical* and *ontological* problem of self-organization which may be informally stated as follows:

> does *self*-organization imply organization of self *by self* or organization of self *by other*[7] ?

This problem is closely related to another within philosophy, viz. that of relating being to becoming. Rescher (1996), maintains that according to Aristotle, "something which is coming into being cannot (ex hypothesi) be said to exist as yet." (p.126) However, he contests this position in arguing that

> at the point of its genesis, a substance is already somehow at hand, since `it' already is doing something - viz., emerging into existence - so that there already is a something to which that coming-into-being can be attributed. We therefore cannot flatly and unqualifiedly deny the thing's existence either. The claim of its existence is thus neither strictly true nor strictly false. (pp.126-127)

There are a number of points to note in connection with the above statement: First, Aristotle and Rescher are committed to different *concepts* of existence. For the former, a substantialist, existence is *static* in the sense that it atemporally (that is, ahistorically) defines the substantiality of a substance (chapter 6). For Rescher, by contrast, existence is *dynamic* (temporal, historical) *because* processual; second, it could be argued that it is only *a posteriori* the emergence of the substance (being, phenomenon) that `it' can be identified *as* a substance (that is, as an `it') and `its' genesis and becoming associated with it; in short, prior to its attaining stability *as* a substance, there is simply is no `it' because the self-other distinction has not yet emerged; third, and consistent with this latter

[7] Here `self' is to be understood as synonymous with thing, object, entity etc, and `other' with context, world, environment etc. *Experiential* connotations of subjectivity and objectivity are not intended.

position, it might be argued that while the substance (`it') is emerg*ing*, it is not `it' itself but `its' *ground*[8] which *brings* `it' into existence. Ontically-speaking, the relationship is that of ground to consequent with the former both causally and temporally prior to the latter. (Inclusion of the temporality condition necessitates excluding the possibility of retroactive causation which is discussed in section 3.6.2.) However, in the above statement, Rescher fails to establish the link between existence and causality, which Elstob (1984) encapsulates in the following maxim, viz. "something exists if it makes a difference". A corollary of this maxim is the necessity of there existing at least two existents (beings) in which case the self-other distinction already holds and hence, is non-emergent (that is, not self-organizing) since each existent (self) must act causally on (at least) one other *in order to exist*. On this view, there cannot, *ultimately*, be any *absolutely* self-organizing *beings* since all *beings* are to some extent other-organized.

However, this does not, thereby, entail the view that absolute self-organization *as such* is impossible. According to Heidegger, what is historically prior to the Aristotelian being-becoming and self-other dualities is *aletheia-physis,* the emerging power which brings itself forth (appears or *emerges*) into presence from concealment. Although this appears to merely restate Rescher's above argument for interpreting a being in the *process* of becoming as a *self-organizing* being (existent), this is, in fact, not the case since *aletheia-physis* does not refer to a being or class of beings (existents) but to *the Being of beings*, that *ground* of which beings partake which has both continuous and discrete (chapter 1) and static and dynamic (chapter 6) aspects and which is itself not a being. (On this basis, Heidegger's (1959) seemingly paradoxical assertion to the effect that Heraclitus - archetypal philosopher of change and process - and Parminides - archetypal philosopher of stasis and substance - are in ontological agreement[9] can be explained by identifying beings and becoming(s) as different aspects of the *manifestation* (or unconcealment) of Being.)

Thus, while Heideggerian thought and process metaphysics both *appear* to provide a means for connecting being and becoming, it is, in fact, only the former which is able to do so. This is because on Heidegger's view (chapter 6), Being is modally-*pluralistic*, manifesting as both substance *and* process, neither of which is ontologically-primordial relative to the other; processualism (chapter 2), by contrast, is modally-*monistic* since substances are held to emerge *from* - and hence, are ontologically-derivative relative to - processes. (This categorially-monistic aspect of process *as such* holds irrespective of whether processes are ultimately reducible to a single processual kind or are type-

---

8 A Heideggerian interpretation of the concept of ground is presented in chapter 6.

9 As Heidegger states, "Heraclitus, to whom is ascribed the doctrine of becoming as diametrically opposed to Parminides' doctrine of being, says the same as Parminides." (p.97) However, as will be shown in chapter 6, this `sameness' does not necessarily entail identity such that dynamic becoming is *reduced* to static being; rather, on Heidegger's view, sameness entails *unitary relatedness*.

pluralistic.) However, this position is problematic since it is unclear how stable entities of any particular kind can emerge in a non-formistic - or, more precisely, purely *immanentist* - processual substrate[10]. Types, kinds, classes, categories etc are fundamental to human discourse (linguistic or otherwise): Humans are `cutters' (that is, categorizers or classifiers) of reality (chapters 1 and 6). Yet if human ontology is processual, how *can* the `cutting' ability emerge since `cutting' implies discretisation while processes are continuous[11] ? The problem therefore reduces to explaining how discreteness arises from continuity and has led some processually-inclined system theorists to adopt a form of symbol-matter dualism as an alternative to a pure process philosophy (Cariani,89).

Additionally, in its recognition of the historicity of the relationship between Being and beings or - ontically-speaking - between ground and consequent, the Heideggerian view provides a more appropriate framework within which to examine notions of self-organization and emergence. For example, returning to the original problem, viz. does self-organization refer to organization *by self* or *by other*: On the Heideggerian view, discrete self-other distinctions imply a duality which is inconsistent with the continuum aspect of the Being-beings relation; however, if the ontological connection of self with

---

10 This is the basis of Collingwood's (1945) criticism of Alexander's (1920 ontology (chapter 5). Whitehead's (1926) solution to the problem is to incorporate a Platonic realm of ideas into process metaphysics, viz. the *eternal objects*. However, Pepper (1942) maintains that this results in an `eclectic [or confusing ontological] mix'. Maturana and Varela's (1980) autopoietic solution will be examined in chapter 6.

11 In the context of a discussion of the emergence of events in perception, Avrahami et al. (1994) present what they refer to as the *cut hypothesis* which has two forms: (1) Form I - "a sub-sequence of stimuli is cut out of a sequence to *become* a cognitive entity *if* it has been experienced many times in different contexts [emphasis added]." (p.239); and (2) Form II - "a sub-sequence of stimuli is cut out of a sequence to become a cognitive entity *for someone*, if it has been experienced many times, with different sub-sequences preceding and following it on the various occasions [emphasis added]." (p.245) Crucially, on their view, "definition of events by goals or by `our concerns' is .. problematic [because] in order to explain how a subject may decompose a temporal sequence, we need a theory of goal structure. Lacking it, the investigator can at best apply her or his own conception of goals, her or his own structures of meaning. Such a method would lack any predictive power whatever." (p.243) There are (at least) two points to note in connection with the latter statement: First, it is quite possible that while specific goals are subjective (that is, particular), goal structure is itself objective (that is, universal). This position is supported by Heidegger's analysis of the existential structure of human being or *Dasein* (chapters 1 and 6) and his distinction between concrete-particulars (existent*iells*) and concrete-universals (existent*ials*), both of which are defined intentionalistically (that is, teleologically) in terms of concerns, *proximal* (or relative) and *distal* (or absolute) respectively. It is interesting to note that, according to Avrahami et al., "although `our concerns' determine the boundaries of what we perceive, the `concerns' themselves derive from perceiving reoccurrence." (p.259) However, this position is problematic since for something to be identified *as re*-occurrent, it must first be identified *as occurrent*, thereby necessitating the existence of an *a priori* - and hence, non-emergent - classification scheme (for example, Kantian schemata). In this connection, it is significant that Avrahami et al. accept the following problem with the cut hypothesis, viz. "what aspects of similarity are necessary for one sequence to be regarded as a recurrence of another. Without an answer to this question, the conditions required for categorization cannot be explored, and the full account of the emergence of events cannot be given." (p.246); second, as will be seen in what follows, it is lack of predictive power which characterizes the emergent event *as* emergent.

other is accepted, it follows that self-organization must be interpreted as organization of self by *that which is prior to self and other* because, ontically-speaking, Being is to being(s) as ground is to consequent(s) as prior is to self and other (chapter 6). Epistemological support for this position is provided by *observational-relativism* (section 3.8.2): The frame of reference of an observer (system-describer) can always be *reduced* such that system is subsumed into environment; thus, it might appear that in the limit there is no system only environment. However, in this event there is no system *or* environment, only that which is prior, viz. the undifferentiated substrate. The role of the observer or *epistemological* `cutter' (chapter 1) is fundamental to the discussion on emergence and self-organization since the observer *defines* the conditions under which Being manifests as beings and things become identifiable *as* things. However, an interesting case to consider is the reverse, when the frame of reference of the observer is *enlarged* such that system and environment become identical; again, it might appear that in the limit there is no environment only system when, in fact, there is only the undifferentiated substrate. Is self-organization possible in these limiting situations ? These issues are fundamental to the debate on emergence within `closed' computational systems and are addressed in sections 3.4.1 and 3.9.1.

Finally, and linked to the semantic problem of self-organization understood from the perspective of the relation between being and becoming, it is unclear whether or not it is meaningful to refer to things which do not physically exist, that is, things which neither are nor have material (physical) referents. Unicorns and the present King of France are examples of such things. Mathematical intuitionists would also include certain types of numbers such as the irrationals whose expansion is infinite since they cannot be constructed in a finite universe: Irrationals *can* be generated algorithmically since finite, terminating procedures exist for producing successive digits in their expansions; however, the expansions must be finite if the universe is itself finite and if such expansions are to *physically* exist (chapter 2). Hence, *if* the universe is finite, such entities clearly cannot be part of the material world. However, entities such as human beings are clearly also capable of conceiving them (at least, in the abstract). How is this *possible* if human beings are finite, physical entities ? There *appear*[12] to be two basic solutions to this problem: Either (i) there exists a Platonic realm of *subsistent* forms to which the mind has access which contains such physically non-realizable entities (chapter 2) or (ii) the mind (human, and possibly other kinds) must somehow be capable of *generating* abstract categories (including those containing non-realizable infinites). If the latter position is adopted and the universe is, in fact, finite and physical then mind must, in some sense, be a non-physical phenomenon. As Rosen (1991) has pointed out, the relation between subsets and sets is not the same as that between sets and `setness'; the last term is a conceptual *abstraction*, and hence, necessarily non-physical. Both possibilities, viz. Platonism and generative (human) capacity for abstraction lead to

[12] In fact, there is a third possibility, grounded in the ontological difference between Being and beings (chapter 6), and involving the *poiēsis* (coming forth) of existential modalities as instances of the *givenness* of Being.

consideration of the notion of *poiēsis* (chapters 1 and 6): Human beings as teleological entities can conceive of entities before the latter are brought forth into material existence. Moreover, the actualization of a conceptual potentiality by an intentional agency involves *artificing* or artifact production in two ontical modalities (potential and actual). To what extent does postulating the existence of a self (being, thing) prior to its coming into being necessitate postulating the existence of an intentionalistic (that is, *a priori* teleological) agency capable of both conceiving (as a potentiality) and realizing (as an actuality) this self and what implications does this have for self-organization as such ? These questions necessitate investigating notions such as artifactual-emergence and other-organization and are addressed in chapters 6 and 7.

3.4. Basic Concepts

In this section various concepts associated with contemporary theories of emergence derived from cybernetics and systems theory are briefly examined.

3.4.1. Systems

The idea of a *system* is a basic concept within modern theories of emergence and a conceptual prerequisite of functionalism (chapter 1) and artificiality[13] (chapter 4).

Def^n: a system S (or internal environment) is (3.1)

S_C a set of components c_p where $p \in P$ (some finite index set)
$S_R : S_C \rightarrow S_C$ a set of relations r_q where $q \in Q$ (some finite index set)
$S : S_C \times S_R$ a network of relations between components

A system may be viewed as performing a *function* if (i) a second set of components and relations defining the *environment* E external to S is specified and (ii) an input-output (coupling) relation between E and S is defined:

Def^n: an environment E (or external environment) is (3.2)

E_C a set of components c_m where $m \in M$ (some index set)
$E_R : E_C \rightarrow E_C$ a set of relations r_n where $n \in N$ (some index set)
$E : E_C \times E_R$ a network of relations between components

13 The link between systems and artificiality is established by Simon (1981) as follows: an artifact is an *interface* between an *inner environment* or system and an *outer environment*, the environment proper of the system.

Def^{n}: an input-output (coupling) relation I_R-O_R is (3.3)

I_R : $E_C \rightarrow S_C$ a set of relations r_i where $i \in I$ (some index set)
O_R : $S_C \rightarrow E_C$ a set of relations r_j where $j \in J$ (some index set)

Thus, the system maps a set of inputs (from the environment to the system) to a set of outputs (from the system to the environment). This input-output mapping *defines* the functional pathways of the system. However, this description is incomplete since only the topology (structure) of the system-environment coupling has been specified. In order to complete the description it is necessary to define each component and each relation: For example, in a cellular automaton (chapters 2 and 5), each component is a finite state machine (FSM) and relations are connections between FSMs. Furthermore, as Heylighen (1993) states, "no absolute distinction can be made between internal and external, i.e. between system and environment. What is `system' for one process is `environment' for another one." (p.3) Hence, the system-environment coupling relations are relativistic irrespective of whether this relativism is functional (ontological) or observational (epistemological). (There is no system-environment boundary in a CA; the system *is* the environment. However, system-environment boundaries can be identified between computationally-dynamic space-time structures supported by the CA although it is unclear whether these boundaries are ontological or epistemological.)

The relations between a system and its environment are shown in Fig 3.1:

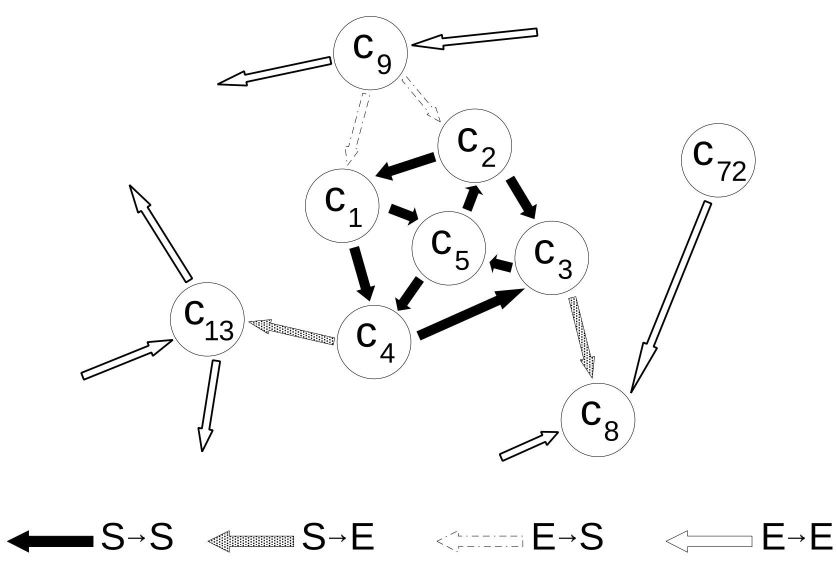

Fig 3.1 Example of a system-environment coupling.

Essential to definitions (3.1)-(3.3) is that *S* be finite while *E* can be finite or infinite: If both *S* and *E* were infinite, it would be impossible to distinguish a system from its environment; *S* would be identical to *E* (and visa-versa). On the above definitions, system and environment are conceptually similar with respect to form. This is consistent with an autopoietic (or self-generating) systems perspective (Maturana,80) in which the connection between an organism and its environment is viewed in terms of a network of perturbation relations between coupled systems[14]. (The ontology of autopoietic systems is examined in more detail in chapter 6.)

The link between systems and emergence may be stated as follows: Emergence occurs during the becoming or coming-into-being of the system-environment duality. A system, as defined above, necessitates description at two levels, viz. (1) the component (or parts) level and (2) the system (whole or functional) level. Furthermore, the coming-into-being of systems fufills one of the requirements for emergence, viz. non-reductionism, since system descriptions cannot be *epistemologically* reduced to descriptions at the component level, even if causality is strictly `bottom-up' (section 3.4.2); components (parts) are only identifiable with respect to systems (wholes) and visa-versa. (The *ontological* possibility of `top-down' or retroactive causality is discussed in section 3.9.2; in this latter case, it is not only description which is two-level but causation as well.) However, this view of the link between systems and emergence is incomplete since it fails to distinguish systemic emergence from chaos, whereby the latter is meant simply the unpredictability that arises as a consequence of the non-linear properties of certain dynamical systems such as Newtonian *n*-body systems (chapter 2). In order for a system to qualify as genuinely emergent, it must possess properties which its components do not possess and these are broadly of two kinds, viz. structural and functional (section 3.5.2).

3.4.2. Hierarchies

The two-level description of systems leads directly to the notion of a *hierarchy*. Simon (1981) defines a hierarchy as

> *a system that is composed of interrelated subsystems, each of the latter being in turn hierarchic in structure until we reach some lowest level of elementary subsystem*. (p.196)

It is important to note, however, that this recursive definition is motivated by purely epistemological concerns; it states how a system may be observed or *interpreted*. Ontological issues, for example, how the system is causally organized or *produced*, are not addressed. Adopting the causal perspective enables two different approaches to hierarchy construction to be identified, viz. top-down and bottom-up.

---

[14] Maturana and Varela (1980) reject the definition of autopoietic systems as stated in conventional systems-theoretical terms (that is, in terms of inputs and outputs) maintaining that the correct characterization of the organism-environment relation is in terms of exosystemic perturbations and endosystemic organizational homeostasis. However, their point is moot since, irrespective of interpretation, such relations do in fact exist.

- *top-down* (designed or planned)

The system subsists as an *a priori* potentiality in the Platonic sense of an abstract transcendent form (structure) or mental artifact (concept) awaiting actualization in matter; on this scheme, the system-environment boundary is pre-defined.

- *bottom-up* (emergent or evolutionary)

The system comes into being as an *a posteriori* teleological actuality, form (structure) and *telos* (function) arising as a consequence of the establishment of a coupling relation between components. Systems become separable and identifiable from their environments once partial or *relative* causal closure is established in the former, assuming this is possible (section 3.9.1).

A detailed account of the conceptual issues associated with hierarchies is given in chapter 5 in the context of a presentation of the unifying framework of computationally emergent artificiality developed and described in this thesis.

3.4.3. Partitionings

System hierarchies must be distinguished from system *partitionings*: In the former, both the set of subsystems and their interrelations are necessarily specified, while in the latter, specification of the subsystems alone is sufficient. Fig 3.2 shows the difference between system hierarchies and system partitionings.

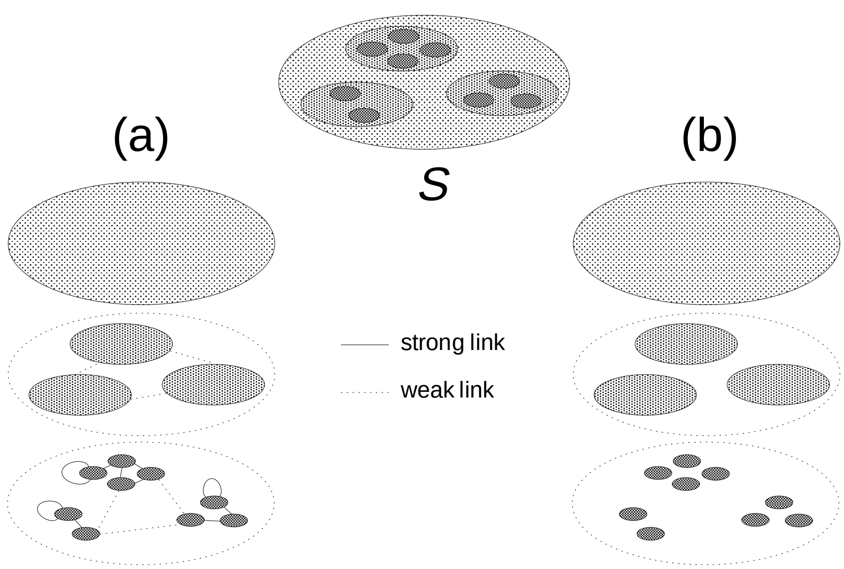

Fig 3.2 (a) system hierarchy (b) system partition.

It should be noted, however, that systemic structures other than hierarchies and partitionings are possible: Hofstadter (1979), for example, describes a system of `tangled loops' or *heterarchies* in which the relations between levels are non-linear and characterized by a high degree of feedback. According to Silberstein (1998), emergent phenomena are characterized by heterarchical relationality, a view which derives support from Heylighen's (1993) assertion that

> all fundamental types of abstract order which can be found in emerging systems, such as hierarchies, symmetries, periodicities, cycles, partitions, etc can be generated by the recursive combination of extremely simple `closure' operations [and] the emergence of cyclical distinctions .. may lead to heterarchical, non-linear architectures. (p.5)

3.4.4. Near-Decomposability

Simon (1973) maintains that multi-level hierarchies constitute the most stable and efficient structures for systems of even moderate *complexity* (section 3.4.5): The time required for a system to evolve by natural selection is much shorter (logarithmic) if the system is structured as a series of levels of subsystems rather than composed from the primitive system components. Simon (1981) further maintains that the bottom-up construction of hierarchies tends to produce systems which are either decomposable or *near-decomposable*, systems in which intra-subsystemic interactions are strong (high cohesion) and inter-subsystemic interactions neglible or weak (low coupling) respectively. Near-decomposability ensures that the short-term behaviour of each of the component subsystems is approximately independent of the short-term behaviour of the other components; in the long term, the behaviour of any one of the components depends in only an aggregate (or statistical) way on the behaviour of other components. Near-decomposability facilitates abstraction and *reductionism* (section 3.8.3) by an appeal to the "empty worlds hypothesis", viz. most things are only weakly connected to most other things; hence, "for a tolerable description of reality only a tiny fraction of all possible interactions needs to be taken into account." (p.221) However, what constitutes a `tolerable' level of description varies depending on context, the level of precision required being motivated by *relevance* to either the observer/interpreter or/and organizer/producer of the system.

- ☐ Historicity

Near-decomposability has been criticised by Wimsatt (1972) who maintains that it fails to distinguish the decomposability or stability of subsystems *before* they aggregate in the system from their decomposability or stability as isolated components *after* they have aggregated. Such criticism appears justified in the context of biological systems assuming (as Wimsatt does) geological time, which allows for a process of mutually coadaptive changes under the optimizing forces of natural selection. This important observation has also been made by Kenyon (1941), who points out that a possible consequence of emergent evolution is that the parts in the new whole may be modified:

For example, a set of components *{a,b,c,d}* might interact to form the emergent whole *X*. When *X* is subsequently fractionated (decomposed) along `natural' hierarchical boundaries, *X* may reduce to components *{e,f,g,h}* where *{a,b,c,d}* is not isomorphic with *{e,f,g,h}*. Thus, emergent evolution redefines the components of the whole via dynamical construction of a hierarchical structure within the whole. This position is consistent with the non-cumulationist[15] account of emergence proposed by Bunge (1969) and discussed in chapter 5.

□ Relativity

Rosen (1977) maintains that because a system can be decomposed on a number of equally `natural' partitionings, objective decomposition is impossible. However, Simon (1981) argues to the contrary by appealing to the role of what may be identified as Kantian synthetic *a priori* (that is, categorial `filters') in observation and perception:

> The fact .. that many complex systems have a nearly decomposable, hierarchic structure is a major factor enabling us to understand, describe and even `see' such systems and their parts. Or perhaps the proposition should be put the other way round. If there are important systems in the world that are complex without being hierarchic, they may to a considerable extent escape our observation and understanding. Analysis of their behaviour would involve such detailed knowledge and calculation of the interactions of their elementary parts that it would be beyond our capacities of memory or computation. (pp.218-219)

Simon reinforces this objectivist conception of system decomposition by maintaining that "in most systems in nature it is somewhat arbitrary as to where we leave off the partitioning and what subsystems we take as elementary." (p.196) However, this view may be contested on the grounds that it fails to recognize the role of the observer-interpreter who identifies a system *as* a system by regarding certain phenomena (properties, behaviour etc) as systemic and others as non-systemic according to certain relevance criteria. Non-systemic phenomena are considered `side-effects' and for certain types of system, for example, linear dynamical systems, may be negligible. However, for other types of system, such as non-linear dynamical systems, side-effects can be the source of an exponentially-increasing disorder leading to increasing *unpredictability* (an epistemological problem) and *instability* (an ontological problem) of system behaviour.

□ Locality

---

15 Silberstein (1998) is also committed to this position, viz. "all ontologically emergent properties/entities have unique causal capacities that constrain or supersede the behaviour of the parts in question. But some cases of emergence present us with even more radical violations of mereological supervenience or part/whole reductionism. [For example, there are cases of `fusion' which] is characterized by the original or subvenient property going out of existence or being `used up' in producing the emergent property or `fused instance' .. In such cases, the subvenient property instances literally no longer exist at the same time as the emergent property instance." (p.474)

Another criticism at a much more fundamental level involves the underlying premise on which the notion of near-decomposability is based, viz. weak *inter*-subsystemic interaction. Bechtel (1986) refers to this as the "localizationalist" perspective and proponents of this view argue that although organization *is* present in a system, it is *inessential* to the explanation of the system. However, Bechtel contests this position maintaining that

> sometimes it is the case that a system performs a task because some part of it performs that task. In such cases a localizationist account is correct .. Moreover, even in cases where the localizationist account turns out to be incorrect, it is sometimes reasonable to begin one's research by trying to find localized components which may perform the task one is trying to explain What one learns about the components of the system may show that no component performs the whole task and may provide critical guidance in developing an explanation that attributes a major role to the integration of the components. (p.34)

3.4.5. Complexity

Ever since the theory of emergent evolution was first articulated, it has been appreciated that the occurrence of emergence is contingent on the complexity of the substrate. For example, Langton (1990,1991b), following a basic insight due to von Neumann (1966), postulates the existence of an `edge of chaos', a region of self-organized criticality lying between ordered (periodic) and random (chaotic) behavioural regimes to which, it is conjectured, certain kinds of complex non-linear dynamical systems evolve. Although complexity is a *necessary* condition for emergence, to infer that complexity is *sufficient* for emergence does not constitute an explanation; the nature of the relationship between complexity and emergence remains to be specified. However, the problem is further compounded by disagreement over the definition of the metric used to measure the complexity of a systemic substrate (component level). Edmonds (1995) describes various complexity metrics in the context of a study of the link between complexity and biological evolution and Bennett (1990) presents the following list of metrics gathered from a review of the literature associated with measuring complexity in physical systems:

- Life-like properties
- Thermodynamic potential
- Computational universality
- Computational space-time complexity
- Algorithmic information
- Long-range order
- Long-range mutual information
- Logical depth
- Thermodynamic depth
- Self-similar structures and chaotic dynamics

Disagreement over how to measure complexity has led, in at least one case, to the view that it may be impossible to provide an objective definition of the concept (Horgan,95). For example, von Neumann argued that complexity was the essential characteristic of life. However, Bennett (1990) has criticised this view on the grounds that it does not differentiate between *potential* and *actual* complexity; the former is a static, structural concept while the latter is a dynamical, control concept (Pattee,73b). On the basis of these arguments, it is proposed that complexity should be regarded as an *observationally-relativistic* measure (section 3.8.2) in the sense that it is defined with respect to epistemic perspective: The emergent properties of a system in which we are interested *define* (via selection amongst alternatives) the metric we use to measure substrate complexity. This position is consistent with that of Tallis (1994) who argues against the idea of *intrinsic* complexity. However, Edmonds (1995) maintains that even relativistic formulations of the concept are problematic on the grounds that notions of complexity can usefully be applied only to constructions *within a given language*; this leads to a consensually-defined as opposed to objectively-defined concept of complexity[16]. Edmonds proposes the following definition of complexity, viz.

> that property of a language expression which makes it difficult to formulate its overall behaviour, even when given almost complete information about its atomic components and their inter-relations. (p.6)

On his view, complexity is observationally-relativistic and "only revealed through interaction with the complexity of another system (typically us)." (p.3) It might be argued, however, that since complexity metrics are products of human design or *anthropo-artifacts* (chapter 6), complexity is a concept which is only meaningful with respect to human artificers (producers); to assert otherwise entails a charge of anthropomorphism, that is, *projection*[17] of human characteristics onto non-human entities. Clearly, this is an issue which impacts heavily on the debate regarding the sufficiency of computationalism as the metaphysical basis for a unifying framework of emergent artificiality.

---

16 This position is consistent with Maturana and Varela's (1980) consensualist interpretation of cognitive linguistic function, that is, autopoietic *languaging*.

17 According to Tallis (1994), *morphization*, projection or "transferred epithet" is usually bi-directional, that is, properties associated with one class of beings are associated with another *and visa-versa*. For example, "machines are described anthropomorphically and, at the same time, the anthropic terms in which they are described undergo a machine-ward shift. These same terms, modified by their life amongst machines, can then be re-applied to minds and the impression created that minds and machines are one." (p.2) In this connection, the following statement by Gould (1986) is highly significant, viz. "if we take the mathematical structures devised to describe the worlds of celestial *mechanics*, statistical *mechanics*, quantum *mechanics*, continuum *mechanics* ... and just plain *mechanics*, and then map the human world unthinkingly onto such structures, is it possible that the human world so decribed can be anything but *mechanical* ? In brief, does the mathematical `language' chosen allow the description, and allow our thinking, to appear as anything but *mechanistic* ?" (p.3)

3.5. The Phenomenology of Emergence

In the following section, the phenomenology[18] of emergence is described. Various concepts of emergence are discussed and different types of emergent are presented.

3.5.1. Concepts of Emergence

Similar to the four concepts (as contrasted with types or kinds) of computation postulated by Emmeche (chapter 2), there are various *concepts* of emergence which can be identified. For example, Pattee (1989) presents the following three concepts:

- *syntactic* emergence (for example, symmetry-breaking and chaos), characterized by non-symbolic, rate-dependent, continuous dynamical systems;

- *semantic* emergence (for example, genetic and cognitive creativity), characterized by symbolic rate-independent, discrete dynamical systems in which it is possible for the emergent to stand for a referent. "Semantic emergence operates on existing data structures that are the result of completed measurements or observations" (p.73);

- *measurement* emergence, characterized by the production of a record of some type of classification of the environment in a measuring device.

Cariani (1989,1991) also presents three concepts of emergence:

- *Computational* (formally based, Platonistic) in which global patterns arise from local micro-deterministic computational interactions; for example, \`gliders' in the Game of Life (chapter 2) or swarming and flocking behaviour in artificial ecologies etc.

- *Thermodynamic* (physically based, materialistic) in which stable structures arise as a consequence of self-organization in physico-chemical systems far from equilibrium; for example, dissipative structures (Prigogine,84). However, Anderson et al. (1987) have argued that the structures emerging in dissipative systems are, in fact, unstable. Simon (1981) postulates a statistical concept of emergence based on the notion of near-decomposability in hierarchical systems.

- *Relative-to-a-model* (functionally based, pragmatic) in which emergence is defined as the deviation of the behaviour of a system from an observer's model of it: if an observed system changes its internal structure and behaviour to such a degree that it becomes necessary to modify the observer's model in order to continue to predict its behaviour, then such systemic changes are considered emergent relative to the model.

Emergence-relative-to-a-model is similar to a concept of emergence due to Nagel (1961)

---

[18] Phenomenology as understood in the context of the present discussion should not be confused with its technical use within philosophy as described in chapters 1 and 6; in the current context, phenomenology simply refers to the identification of a concept of emergence and its associated emergents (products), that is, a metaphysics and the phenomena which follow. More precisely, a phenomenology is defined by a *concept* of emergence and specification of *types* of emergent.

who maintains that "although a property may indeed be an emergent trait relative to some given theory, it need not be emergent relative to some different theory" (pp.370-371). On this view, "the distinction between an emergent trait and a nonemergent one would shift with changes in interest and with the purposes of an inquiry." (p.377) Umerez et al. (1993) differentiate three forms of the concept of emergence-relative-to-a-model:

(1) emergence with respect to a concrete[19] model: *error*

(2) emergence with respect to a concrete theory: *change*

(3) emergence with respect to a concrete paradigm: *crisis*

(1) necessitates a change in the model, but within a fixed theoretical framework; (2) necessitates a change in both the model and the theoretical framework, but within the scope of the existing paradigm; (3) necessitates a complete change of the entire approach (model, theory and paradigm). Umerez et al. associate (2) with *self-organizing* behaviour in which the relations between the observables at the two levels (substrate and emergent) are fixed or static; this type of self-organization is described as structural emergence. (3) is associated with *functional* behaviour and the relations between the observables at the two levels (substrate and emergent) are dynamic.

The most popular concept of emergence within the philosophy of mind is that due to Churchland (1985) in which emergence is defined as a *relation* between a phenomenon and two theories describing it, viz.

> a property *P* specified by its embedding theory *T*1 is emergent with respect to the properties of an ostensibly reducing theory *T*2 just in case
>
> 1. *P* has real instances;
> 2. *P* is co-occurrent with some property or complex feature recognized in *T*2, but nevertheless
> 3. *P* cannot be reduced to any property postulated or definable within *T*2.

Manthey (1990) supports this view maintaining that "an emergent phenomenon is a product of a hierarchy relation between conceptual levels". What is problematic about this view is that it is unclear whether the emergent relation is objective or subjective; in the latter case, emergence depends on the choice of observation frames adopted by the theoretician and Churchland's concept reduces to the second form of emergence-relative-to-a-model described by Umerez et al. (1993).

3.5.2. Types of Emergent

There are a number of different types or kinds of *emergent* reported in the literature; however, many of them are associated with specific concepts of emergence. Emergents

[19] A *concrete* model is a model of a physical system.

are properties (or qualities) arising during a process and fulfilling one of the following conditions (Klee,84):

Property *P* is emergent at a level of organization in a system, with respect to the system's lower-level microstructure *MS*, when and possibly only when either:

(1) *P* is unpredictable in principle from *MS* (i.e. unpredictable even from an ideally complete theoretical knowledge of *MS* in the limit of scientific inquiry)

or

(2) *P* is novel with respect to *MS*

or

(3) *MS* exhibits a much greater degree of variance and fluctuation from moment to moment than does the level of organization where *P* occurs; *P*'s constant and enduring presence in the system would not seem to be wholly determined by *MS*

or

(4) *P* has direct determinative influence and effects on at least some of the properties in *MS*.

Interestingly, and as Klee states, (3) is consistent with supervenience (section 3.2.4) and functionalism, viz. "multiple distinct micro-bases for the *same* macro-property" (p.56). More important, however, is the fact that (i)-(iii) are consistent with a cumulationist view of emergence while (iv) necessitates adopting a non-cumulationist position (chapter 5).

Emergent properties (or, more simply, emergents) broadly fall into the following two categories:

- *structural*

In the context of a discussion of emergence occurring in systems of distributed, concurrent and asynchronously-coupled computational processes, Manthey (1990) defines the necessary and sufficient conditions for structural emergence[20] as follows:

Necessary conditions

1. There are at least two processes
2. These processes interact via either memory or synchronization
3. The putative emergent phenomenon [invariant structure] cannot - even in principle - be expressed by a single process.

Sufficient Conditions

1. The phenomenon occurs.

---

[20] In this context, a structural emergent is simply a stable configuration in space-time, irrespective of whether this configuration is static or dynamic (periodic).

However, Manthey admits that the "delicate issue of observation" is ignored on this scheme.

- *functional*

Heylighen (1993) identifies at least three kinds of functional emergents, viz. boundaries, organizations, and control relations. However, it is an open issue whether such functions are intrinsic (ontological) or extrinsic (epistemological); in the latter case, emergence is relative to an observer. While it is certainly the case that something must have a function in order to be considered a system[21] (section 3.4.1), it does not follow that systematicity is necessarily an intrinsic property of certain configurations of matter. Even if it is held, following Pepper's conjecture (1926), that there can be no emergent properties other than emergent `laws', there is still the problem that a law is defined in terms of a set of observables which necessitates the existence of an observer. This leads directly to the epistemological problem of intrinsic and extrinsic emergence discussed in section 3.8.1.

3.6. Formalism, Computationalism and Emergence

In this section, various formalisations of the concept of emergence are briefly examined in preparation for the presentation in chapter 5 of a unifying framework of computationally emergent artificiality. Simon (1981) anticipates a formal framework for emergence in his conception of hierarchical systems, viz.

> hierarchic systems are usually composed of only a few different kinds of subsystems in various combinations and arrangements .. Hence we can construct our description from a restricted alphabet of *elementary* terms corresponding to the basic set of elementary subsystems from which the complex system is generated. (p.221)

3.6.1. Basics

In order to explain the behaviour of any whole in terms of its structure (or organisation) and components, the following two independent pieces of information are necessary (Broad,25):

- how the parts behave separately, i.e. when not in the whole
- the law(s) according to which the behaviour of the separate parts is compounded when they are acting together in any proportion and arrangement

Such requirements translate isomorphically into a distributed formal-computational

[21] Contrary to the assertions of Maturana and Varela (1980), this holds for autopoietic as well as non-autopoietic (or allopoietic) systems since the former have organizational-homeostasis as their function. Hence, it is argued, functionality is essential to systemicity.

model of the corresponding system (Langton,89b), viz.

- a description of the behaviour of each member of a set of *behavors*, each of which is an automaton (for example, finite state machine (FSM), Turing machine (TM) etc)
- a description of the set of local *rules* governing the interactions between behavors (for example, the state-transition rule for the FSM, Turing machine specification etc)

3.6.2. A Formal Framework for Emergence

Baas (1993,1997) describes a formal framework for emergence involving three components which are arbitrarily defined, viz. primitives, observations, interactions:

Primitives $\{S_i^1\}$ is a set of first-order structures, where $i \in J$ (some index set, finite or infinite).

Observations $Obs^1(S_i^1)$ denotes the properties of structure S_i^1 registered by the observational mechanism Obs^1 which may be internal or external to the system.

Interactions Int^1 is the set of interactions between the elements of $\{S_i^1\}$ allowed by $Obs^1(S_i^1)$.

$S^2 = R(S^1{}_i, Obs^1, Int^1)_{i\in J}$, where R is the result of interactions between primitives and S^2 is a second-order structure.

First-Order Emergence P is an emergent property of S^2 iff $P \in Obs^2(S^2)$, but $P \notin Obs^2(S^1{}_i)\ \forall\ i$.

The validity of the above framework might be contested on the grounds that it assumes properties to be objectively definable *a priori* the interactions between the objects with which they are associated. For example, on Elstob's (1984) view (section 3.9.1), properties arise during interactions and are contextually-determined *by* the interaction; hence, they are only determinable *a posteriori* the interaction[22].

Baas differentiates between concepts of emergence on the basis of computability or decidability:

- *deducible* (or *computable* or *decidable*) *emergence* - in which there is a deductive or computational process D such that $P \in Obs^2(S^2)$ can be determined by D from $(S^1{}_{i1}, Obs^1, Int^1)$. *Obs* is algorithmically defined.
- *observational* (or *undecidable*) *emergence* - in which P is an emergent property which is non-deducible. *Obs* is non-algorithmic, e.g. the truth function in formal systems capable of supporting arithmetic; such systems have statements whose truth value cannot be deduced within the system (Gödel incompleteness theorem).

Deductive emergence in formal or computational systems is epistemically relative,

22 However, as shown in chapter 6, this position is itself problematic.

describing the extent to which the `Platonic landscape' defined by the attractor-space (or basin of attraction field) of the system has been `mapped' by an observer (chapter 2). (This holds only for finite computational universes or infinite computational universes which are algorithmically compressible.) Simon (1981) alludes to deductive emergence in asserting that "all mathematics exhibits in its conclusions what is already implicit in its premises ... Hence all mathematical derivation can be viewed simply as change in representation, making evident what was previously true but obscure." (p.153) Furthermore, the non-algorithmic nature of *Obs* - in observational or undecidable emergence - can be effectively (practically) overcome using Myhill's theorem as shown in chapter 5; hence, on the computationalist view, it is conceivable that observational emergence might reduce to deducible emergence.

3.6.3. Computational Emergence

Broad (1925) maintains that the difference between mechanism and emergentism lies in the fact that emergence does not occur through substituting certain determinate values for determinate variables in a general law which connects the properties of any whole with those of its separate constituents and with its structure. However, the distinction between emergentism and mechanism is obscured in many computationalist theories of emergence. For example, Darley (1994) presents a rigorous definition of emergence based on the concept of computational complexity (section 3.4.5). On this scheme, emergence is viewed as the result of a phase change in the amount of computation necessary for optimal prediction of certain phenomena, viz.

> emergent phenomena are those for which the amount of computation necessary for prediction *from an optimal set of rules, classifications and analysis, even derived from an idealised perfect understanding, can never improve upon the amount of computation necessary to simulate the system directly* from our knowledge of the rules of its interactions. (p.412)

A similar position, viz. simulation as the optimal means of prediction in emergent systems, is argued by Rasmussen et al. (1995). Darley views emergence in computational systems of finite size and finite time as analogous to undecidability in systems of infinite size and time capable of supporting universal computation. Furthermore, whether a system is emergent or non-emergent is an undecidable proposition. Hence, the term `emergence' may be regarded as an epistemological statement on the undecidability of certain classes of computational phenomena. However, this view fails to distinguish emergence from mere chaos (section 3.4.1) and, more importantly, is contested by Cariani (1989) who maintains that

> for Turing machines with finite tapes, the .. Halting Problem disappears completely, *along with all other computability issues* [emphasis added] (p.178)

the implication being that emergence in finite computational systems - that is, finite state machines - is a decidable issue, viz. such systems are, ultimately, *closed* or bounded in their potential for emergence (Ali,98a).

3.6.4. Emergent Computation and Computational Emergence

As stated in section 3.5.1, computational emergence necessitates a computational substrate. By contrast, *emergent computation* (Forrest,90) necessitates a computational ontology for both substrate and emergent, that is, the phenomenon emerging from the computational process is itself a computational product; hence, emergent computations constitute a subclass of computationally emergent phenomena. Of course, it is trivially the case that the products of computational emergence will themselves be computational since it is assumed that no new ontological kinds are generated in the process. However, Forrest's position is much stronger, being motivated by ideas such as virtual machines and embedded computer hierarchies (chapters 2 and 5), viz. an emergent computation refers to the production either of (i) *components* which can be used in the construction of a virtual machine or (ii) the virtual machine *system* itself. The concept of emergent computation can be understood at three levels:

- *substrate* a collection of agents (or behavors), each following explicit instructions
- *emergent* interactions among the agents (according to the instructions) giving rise to global patterns of behaviour at the macroscopic level, i.e. epiphenomena
- *interpretive* natural interpretation of the epiphenomena in computational terms

This concept of emergence is consistent with the notion of global constraints in systems with computational substrates described in section 3.9.2. Even if the question of the epiphenomenality or otherwise of emergents is ignored, interpreting the epiphenomena in computational terms is a contentious issue because there are at least two possibilities with respect to any such interpretation, viz. (i) computation "in the eye of the beholder" (extrinsic emergence) and (ii) computation inherent in the system itself (intrinsic emergence) (section 3.8.1).

Additionally, it is conceivable that non-computational substrates might give rise to phenomena which are interpretable as emergent computations; in fact, physical computers can themselves be viewed in this way: Computers are material artifacts emerging as a consequence of human design activity. Ontologically, their *operation* is governed by physical laws; however, epistemologically, their *behaviour* can be given a functional interpretation. Temporarily ignoring the possibility of a computationalist interpretation of the laws of physics (chapters 2, 4 and 5), this consititutes at least one instance of the emergence of computation from a non-computational substrate; in this case, the efficient cause of the phenomenon - the means by which the computer comes into existence - is a human artificer (producer, organizer). If the various concepts of computation identified by Emmeche (chapter 2) are accepted, it may be possible to extend the list of computational emergents arising from non-computational substrates. This gives rise to the notion of *heterogeneous* emergent computation wherein a substrate with one kind (concept) of computational ontology facilitates the emergence of a global phenomenon of a different kind of computational ontology; accepting the validity of

Emmeche's scheme, this may be the case in the natural world. However, as Emmeche has argued, such a position ultimately rests on the pluralistic assumption that attributing computational properties to natural phenomena (as contrasted with the computationalism intrinsic to computational artifacts, viz. computers) is metaphysically correct.

Finally, emergent computation may also be interpreted as `environmental' computation since in systems capable of supporting such emergence, it may be argued that information processing occurs by the components in the system (the `organisms' or local systemic level) acting on the system as a whole (the `environment' or global systemic level) (Millonas,94).

3.6.5. Emergence in Cellular Automata (CAs)

There are two problems concerning cellular automata or CAs (chapters 2 and 5) involving issues related to emergence and artificiality: (i) the forward problem and (ii) the inverse problem (Gutowitz,90). The *forward* problem can be described as follows: Given a CA rule, determine (predict) its properties. Clearly, this involves consideration of the epistemological issues associated with emergence; however, it is important to realize that the forward problem is not identical to the emergence problem since it does not suffice to differentiate emergence from unpredictability (chaos). The *inverse* problem can be described as follows: Given a description of some properties, find a rule or set of rules which have these properties. If the properties belong to a natural phenomenon, for example, a living or cognizing organism[23] (assuming they can be listed), then the attempt to realize these properties in a CA assumes (of necessity) the validity of functionalism and multiple-realizability (chapter 2), the basis of artificiality (chapter 4). The phenomenon of emergence in CAs will be examined further in chapter 5 in connection with the presentation of a unifying framework of computationally emergent artificiality.

3.7. Philosophical Issues: Introduction

The problem with Mayr's definition of emergence (section 3.2.2) is that the notion of an emergent `characteristic' is ambiguous. It is possible to refine the definition by adopting the position of Searle (1992) who, following Pepper (1926), distinguishes two kinds of emergence, viz. (i) "Emergent1" (or *property-emergence*) and (ii) "Emergent2" (or *law-emergence*).

3.7.1. Property-Emergence

Emergent1 features (or characteristics) are ontically but not epistemically reducible (section 3.8.3); while their existence may be explained by the causal interactions of

[23] For Maturana and Varela (1980), these terms are synonymous; however, in this study, they are taken to be distinct.

components at lower levels in a system, the *properties* of the emergents are not deducible from the properties and laws governing the interactions of the components. This is consistent with the conventional account of the theory of emergence (or `emergentism') as a materialistic philosophy (Mayr,82).

3.7.2. Law-Emergence

Emergent2 features are neither ontically nor epistemically reducible and are, according to Searle, based on a "much more adventurous conception": A feature *F* is emergent2 if and only if *F* is emergent1 and *F* has causality that cannot be explained by the causal interactions of its components. Searle rejects the possibility of emergent2 on the grounds that the existence of any such features would seem to violate even the weakest principle of the transitivity of causation (Searle,92). However, this view rests on the assumption of universal determinism (chapter 2), a controversial metaphysical position given the apparent indeterminism of physical systems at the quantum level (chapter 4). It appears that Searle (and others) regard emergent2 frameworks as dualistic (mentalistic or vitalistic) positions which introduce into the explanation of phenomena elements which are logically and scientifically unecessary. However, as Wilkinson (1979) has pointed out, the notion of emergent *laws* (as contrasted with emergent properties) need not violate the existing framework of physical laws; emergent laws may complement physical laws by augmenting them, thereby *extending* the set of natural laws[24]. On this view, a phenomenon at a particular level in the (natural) phenomenal hierarchy is ontically-consistent-with but not ontically-reducible-to phenomena at lower levels in the hierarchy[25].

Searle's scheme leads to a consideration of the concept of emergence from two philosophical perspectives:

- *Epistemological* - concerned with issues of interpretation and observation: For example, are phenomena intrinsically or extrinsically emergent ? Is emergence observationally-relativistic or absolute ? Is reductionism possible ?

- *Ontological* - concerned with issues of production and organization: For example, is emergence closed (bounded) or open (unbounded) ? Are emergents causal or non-causal (epiphenomenal) ? Is emergence ontically-monistic or pluralistic, cumulationistic or non-cumulationistic ?

---

24 The notion of emergent or evolutionary laws has been described in (Lovejoy,27) and a recent application of this idea in the context of an emergentist framework resolving inconsistencies between classical (special-relativistic) and quantum (non-local) accounts of physics is presented by Silberstein (1998).

25 As will be seen in chapter 5, this constitutes an instance of ontological-cumulationism.

3.8. Epistemological Issues

In the following sections a number of epistemological issues associated with the concept of emergence including intrinsic and extrinsic emergence, the problem of observational relativism and the limits of reductionism are examined.

3.8.1. Intrinsic and Extrinsic Emergence

For Crutchfield (1994), the appearance of *novelty* in dynamical systems, identified as the phenomenon of emergence, necessitates answering the following related questions: "For whom has the emergence occurred ?" and "to whom are the emergent features `new' ?" (p.2). This leads directly to consideration of the concept of *observation*. Crutchfield identifies two possibilities for the observer of an emergent phenomenon with respect to its *situatedness*, viz.

- *extrinsic emergence*

the observer of the phenomenon is *external* to the phenomenon, that is, the phenomenon is *observed*. Functionality is *contingent* on the assignment of systemic status by the external observer; systematicity is an interpretation.

- *intrinsic emergence*

the observer of the phenomenon is *internal* to the phenomenon, that is, the phenomenon is *self-observing*. Functionality is *necessary* since the phenomenon is performing a function (observation); the phenomenon *is* a system rather than being merely interpretable *as* a system.

Crutchfield interprets emergence in a manner consistent with Cariani's concept of emergence-relative-to-a-model (section 3.5.1), that is, as a non-transient deviation from a normative model of the phenomenon. Intrinsic emergence is described by Crutchfield as follows:

> In the emergence of coordinated behaviour .. there is a closure in which the patterns that emerge are important within the system .. [Such] patterns take on their `newness' with respect to other structures in the underlying system. Since there is no external referent for novelty or pattern, we can refer to this process as `intrinsic' emergence.
>
> What is distinctive about intrinsic emergence is that the patterns formed confer additional functionality which supports global information processing .. during intrinsic emergence there is an increase in intrinsic computational capability (p.3)

Hence, the connection between intrinsic emergence and emergent computation (section 3.6.4). Crutchfield further states that

> the closure of `newness' evaluation pushes the observer inside the system. This requires in turn that intrinsic emergence be defined in terms of the `models' embedded inside the observer. The observer in this view is a subprocess of the entire system. (pp.3-4)

On this view, which is similar to the concept of organizational-homeostasis in autopoietic systems theory (Maturana,80), emergence is primarily conceived as ontological, viz. increased system functionality due to component level interactions; epistemological issues such as novelty to an observer are of only secondary significance. Emmeche (1992) supports this view, viz. "emergent properties must be observable, but they appear because of the system of interactions among the lower-level objects (and not because of observation)." (p.93) Thus, intrinsic emergence assumes a realist position in contrast to the observational relativism (section 3.8.2) associated with extrinsic emergence.

It is important to note that Crutchfield acknowledges use of the concept of a model "in a sense that is somewhat more generous than found in daily scientific practice."(p.4). On his computational-realist or *computationalist* (chapter 2) view, models are held to exist implicitly in the dynamics and behaviour of processes as behavioural entities necessitating "excavation" rather than as cognitive entities "in the eye of a beholder". This position has a number of corollaries:

1. Maintaining that a model of a phenomenon is `in' the phenomenon necessitates holding that both models (artifacts) and natural phenomena are ontically systemic. This position may be termed *systemic realism* and is the metaphysical view which asserts that an isomorphism exists between epistemology (knowledge, representation) and ontology (being, reality).

2. Furthermore, maintaining that a model of a phenomenon is `in' the phenomenon entails tacit *a priori* adoption of an objectivist and exclusivist (that is, non-pluralist) position; on such a view, only one *system* model is possible, viz. that in which epistemology coincides with ontology.

3. However, it could be argued that `excavation' is not an objective process since there are, in fact, many possible models of a system, each capable of being produced from a perspective at least partially determined by concerns intrinsic to the *modeller* (Rosen,77).

4. Thus, it could be maintained that intrinsic emergence ultimately degenerates into extrinsic (or observationally-relativistic) emergence. Hence, emergence-relative-to-a-model or rather relative-to-a-*modeller* (since excavation is performed by an observer external to the system) appears to be the appropriate framework for the concept[26].

---

[26] It could be argued that (4) is a simple *non sequitur* which arises from the conflation of epistemology with ontology and that genuine intrinsic (or observer-independent) emergence is possible. This is because, following

Crutchfield defines intrinsic emergence in terms of three concepts: (i) dynamical systems theory, (ii) computation theory, and (iii) inductive inference, the latter involving algorithmic reconstruction of a minimal machine (formal-computational description) belonging to an assumed model class from a set of primary and secondary observations. These are observations of the phenomenon and detection of regularities in a series of increasingly-accurate models belonging to different model classes respectively. Organization is *extracted* from phenomena using "minimally-biased discovery procedures" (p.2). However, what constitutes a `minimal' set of procedures is an open issue and depends on the assumptions made during the modelling process. Furthermore, induction as a scientific method has been subject to intense philosophical criticism since (at least) Hume[27]. Crutchfield (1994) maintains, however, that

> `emergence' is meaningless unless it is defined within the context of processes themselves .. emergence defined without this closure leads to an infinite regress of observers detecting patterns ... The regress must be folded into the system, it must be immanent in the dynamics. (pp.9-10)

In support of a position lying somewhere between intrinsic and extrinsic emergence, Emmeche (1992) argues for an observational interpretation of selection pressure, viz. "the environment acts as an observer that `sees' and `acts upon' higher-level properties, thereby establishing recurrent interactions within and between the different levels." (p.93) This requires a more sophisticated scheme than simple self-organizing dynamics, viz. differential reproduction of a set of variant self-representations leading to emergent evolution via natural selection. However, while the observer is certainly external to the system, it is not an observer in any conventional sense of the term. Moreover, in the limit of the system-environment boundary reducing such that system and environment become identical (with the system *enlarging* to the environment), this form of emergence transforms into intrinsic emergence.

The intrinsic versus extrinsic argument can, ultimately, be reformulated as a debate regarding claims for two different *truth* schemes, intrinsic emergentists supporting a *coherence* theory of truth, extrinsic emergentists supporting a *correspondence* theory of truth[28]. A similar argument arises in the debate over computationalism, viz. whether it

---

Kant, it is consistent to maintain that multiple models of a phenomenon are possible while simultaneously asserting that the phenomenon (in this case, emergence) *as it is in itself* is unitary; epistemic pluralism is consistent with ontic monism (Searle,95). (A yet more radical position, viz. ontological pluralism (Dreyfus,91), is investigated in chapter 6.) However, this position necessitates rejecting the proposed isomorphism between ontology and epistemology, a relation which is fundamental to the systemic-realist position adopted by Crutchfield and definitive of intrinsic emergence.

27 Chalmers (1982) provides a precise and concise summary of the main arguments against induction.

28 *The Oxford Companion to Philosophy* (1995) defines the *coherence* theory of truth as "a theory of truth according to which a statement is true if it `coheres' with other statements - false if it does not." Crucially, for computationalism, "the theory is more plausible for axiomatic systems where `coherence' can take the definite

is intrinsic or extrinsic (chapter 2).

3.8.2. Observational-Relativism

In section 3.4.5, it was maintained that the phenomenon of emergence depends on complexity: Simple phenomena involve simple linear dynamical interactions between components and the behaviour of such systems is predictable via theory. Complex phenomena involve complex non-linear dynamical interactions between components and the behaviour of such systems is unpredictable even if the principles (laws, rules etc) governing the interactions are deterministic. In such cases, there appears to be no alternative to simulation (Rasmussen,95). However, it has also been shown that the complexity metric can be defined in a number of different, observationally-relativistic ways and, crucially, it is this relativism which provides support for extrinsic emergence. For example, Tallis (1994) adopts a similar position vis-a-vis complexity to that held by Cariani (1989,1991), viz. observational-relativism, in stating that "complexity is in the eye of the beholder". Kampis (1991) clarifies this position by postulating a hierarchical view of system description (section 3.4.2): Any `complex' system composed of a number of components may be transformed into a `simple' system by identifying it as a component of a `complex' system at a higher level of abstraction. Since emergence depends on complexity, complexity on degree of abstraction, and abstraction on the concerns of the observer which are relative, emergence must, therefore, be relative to the observer.

In section 3.4.2, it was argued that emergent processes generate hierarchical structures via bottom-up composition. The reverse process, top-down decomposition or *reduction* of phenomena, will now be examined in preparation for a discussion of the ontological issues associated with emergence in section 3.9.

3.8.3. Emergence and Reductionism

Reductionism, the attempt to explain phenomena in terms of the properties and laws associated with lower levels in a phenomenal hierarchy, is one of the basic tenets of modern science and exists in three distinct forms (Ayala,87):

- *ontological*: This refers to the position that the physico-chemical laws are universal; on this view, biological and psychological phenomena, such as life and consciousness respectively, must not contradict these laws.

- *methodological*: This refers to the position according to which we should attempt to explain all

form of being derivable from the axioms." The *correspondence* theory of truth "maintains that the truth of a proposition *p* requires the following two conditions to be met: (1) it is a fact that *p*, and (2) the proposition corresponds to that fact" and is essentially, a theory of the relation (ismorphism) between representation and reality. A recent defense of the correspondence theory of truth against the pragmatist (or instrumentalist) theory of truth is presented in (Searle,95).

> phenomena from the known natural laws, i.e. one should follow the reductive method whenever possible (Occam's Razor) and to the extent that its application is meaningful. Thus, methodological reductionism is a heuristic position.

- *epistemological*: This refers to the statement that all phenomena can be factually reduced, at least `in principle', to lower phenomenal levels and ultimately to physics and the properties of matter.

Searle (1992) further distinguishes ontological and epistemological reductionism as shown in the following diagram (Fig 3.3):

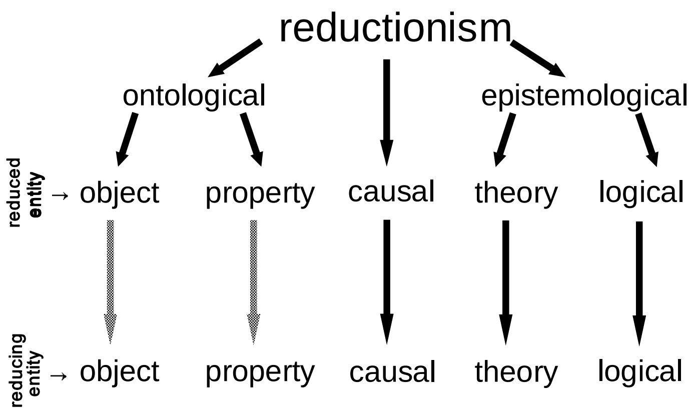

Fig 3.3 Types of reductionism.

At this point in the discussion, it is necessary to determine the relationship between reductionism and emergence. Obviously if the emergence of laws (section 3.7.2) is possible, epistemological reductionism is impossible and methodological reductionism becomes a necessarily incomplete explanatory strategy. In the remainder of this section, discussion will be restricted to the relation between epistemological reductionism and the emergence of *properties* (section 3.7.1) as opposed to laws.

In those cases in which knowledge of the components of a system precedes knowledge of their organised aggregates (for example, in machines and other artifacts) the properties that are said to `emerge' as a result of the act of combination may be regarded as the *realization* of the *potential* of the individual parts (Simon,71). This view is supported by Harris (1965) who maintains that:

> [the emergence of new properties or qualities and levels of integration] is not ineluctable. It is not, of course, always explicable as soon as it is detected, but the novelty is attributable to the new form of organization and once that is fully understood it should be scientifically comprehensible. Moreover, the novelty is a continuous outgrowth from the foregoing phase. It is not a discrete supervenient addition in principle unaccountable in terms of its contributory factors. Certainly, it is not rendered intelligible by analysis alone. It is a product of synthesis and can be explained only by the principles of synthesis. *Potentiality in the earlier and more primitive* of the latter and more

> elaborate must, however, be preserved and the aim of science is always to trace the continuity between them in as great detail as possible [emphasis added]. (p.487)

However, Ayala (1985a) considers this position to be vacuous because it assumes an *a priori* definitional position. To illustrate the point, he cites an argument due to Broad (1925): Are the properties of common salt, that is, sodium chloride, simply the properties of the elements sodium and chlorine when they are combined according to the formula NaCl ? If among the properties of sodium and chlorine, we include their combination in table salt and the properties of the latter, the answer would be in the affirmative (*a posteriori*); otherwise it would be in the negative (*a priori*). The solution is, therefore, definition-dependent, reflecting the epistemic status of the *observer*. Consequently, he maintains that the notion of emergent properties is `spurious', arguing for its reformulation in terms of propositions expressing human knowledge, a position which is consistent with Cariani (1989), viz. emergence-relative-to-a-model (section 3.5.1). As Ayala states,

> the proper way of formulating questions about the relationship between complex systems and their components is by asking whether the properties of complex systems can be inferred from knowledge of the properties that their components have *in isolation*. The issue of emergence cannot be settled by discussions about the nature of things or their properties, but is resolvable by reference to our knowledge of these objects [emphasis added]. (p.75)

On such a view, a property that is regarded as emergent at a particular time might be considered non-emergent at a later more advanced state of knowledge when and if the emergent properties are included in the list of properties associated with the components whose interaction gave rise to the property. Thus, designation of a property as emergent involves the assignment of an epistemic interpretation to a phenomenon by the observer of the phenomenon (Kampis,91). However, there is a problem with the above view with respect to the possibility of component isolation: According to the definition in section 3.4.1, a system exists within an environment unless it is closed, that is, autonomous, in which case system and environment are identical. Assuming that system and environment are not identical, how can a system be isolated from its environment given that its continued existence is predicated on the existence of this environment ? Given the assumption that isolation is possible by `removing' the system from one `suitable' environment (for example, air) and `placing' it in another (for example, a vacuum), it remains the case that what counts *as* the environment for isolation and, as a corollary, the properties of the isolated system depend on the environment selected. Are the `isolated' properties of salt its properties in air or in a vacuum ? Obviously, this is established by a convention which is in turn motivated by a concern for reducing interactions between the system under consideration and other systems. But whether this is true for all systems is an open issue: For example, physico-chemical systems are examples of what might be called *relatively-closed systems*[29]. Vital (or biological) and mental (or cognitive) systems, by contrast, are instances of what might be called *relatively-open systems* (section 3.9.1) for which isolation may be impossible[30].

Theories of emergence require `complete knowledge', that is, knowledge of (i) *all* the properties the components (parts) possess when individually isolated from the candidate system (whole), (ii) *all* the properties possessed by systems which are formed when some or all of the components stand to each other (or to additional components) in relations other than those between the components in the candidate system, and (iii) *all* the properties of the components in these other systems (Nagel,61). It is assumed that the isolation process (that is, the *act* of observation or measurement) generates *objective* information, viz. *the* properties of each component. However, this is only true if the system is uniquely decomposable (section 3.4.4), that is, if *only* one hierarchical description is possible[31]. Simon (1981) argues in favour of system decomposability on the grounds that our epistemological apparatus (sensori-motor system and nervous system including the brain) has evolved to facilitate our `seeing' systems. As stated previously, Rosen (1977) has contested this objectivist position, arguing that a system may be decomposed in many ways motivated by and depending on the interests (intentionality) of the analyst-modeller. Observation and measurement have become for human beings activities which have transcended the possibilities afforded by the body; yet, the *way* in which *we* construct such instruments is at least partially determined by *what we* select for study. Thus, system description is a relativistic activity, and since a component may also be regarded as a system (section 3.8.2), component property description is also relativistic. An attempt at formulating a weaker version of the requirement for isolation could be made by stating the conditions under which isolation takes place; however, the logical conclusion of this is a vicious infinite regress because it is then required to state the conditions under which those isolating conditions hold etc. The result is that isolation cannot be performed objectively, a result indicated by Rosen

---

29 `Relatively' because even in Newtonian mechanics where the effects of physical forces decay exponentially, such effects are still present. Systems separated at great distances are *weakly* connected, thereby allowing for relative system closure and hence, isolation. It should be noted here that this argument assumes an *absolute* view of space and time which, according to special relativity theory (chapter 4), is inaccessible to an endophysical observer, that is, an observer situated within the physical universe.

30 As will be seen in chapter 6, factical science (for example, physics, biology, psychology etc) represents merely one *way* (or modality) of describing the ontology (Being) of phenomena (beings). According to Heidegger, human beings - and, derivatively, *all* other beings - have a way of way of Being, viz. being-in-the-world (or *Dasein*), that is *essentially contextual* in which case, the (Cartesian) approach of atomistically isolating human being in order to determine its essence is incorrect.

31 Whether such a description is attainable constitutes a separate issue concerned with the tractability of the analysis.

in his arguments against objectivism in systems science[32].

3.9. Ontological Issues

Ontology is the study of being and, as metaphysics, of that which is primordial and encompassing (chapters 1 and 6). Consequently, *ontological* inquiry must concern itself at least in part with questioning the validity of the metaphysical assumptions underlying scientific and philosophical frameworks (chapter 1). In the following sections, three ontological issues associated with the concept of emergence are examined: (i) open (unbounded) *vs.* closed (bounded) emergence, (ii) causal *vs.* non-causal (epiphenomenal) emergents, and (iii) pluralistic *vs.* monistic emergence[33].

3.9.1. Open and Closed Emergence

Theories of emergence involve at least two levels of description, viz. the component, local or *substrate* level and the system, global or *emergent* level. On a systems-theoretical emergentist view, a system is held to arise from within a substrate once a set of causal relations attains partial or relative closure (section 3.4.2); as a corolary of such a view, the environment is held to be that part of the substrate which does not form part of the system. Thus, *both* systems *and* environments are emergents[34] (section 3.3). While closure is most usually discussed in the context of systems, it is interesting to examine the concept in the context of its application to environments. There are essentially two kinds of environment (open and closed) which may be differentiated on the basis of two binary parameters, finitude and dynamicity: (i) a *closed* environment (*CE*) is either [1] finite and *static* (both components and relations are defined as fixed sets) or [2] infinite and static; (ii) an *open* environment (*OE*) is either [1] finite and *dynamic* (components and/or relations are defined as variable sets) or [2] infinite and dynamic. The various forms of open and closed environments are shown in Fig 3.4 (overleaf).

---

[32] This position is consistent with arguments establishing the validity of the Frame Problem (chapter 7) in artificial intelligence.

[33] The related issue of cumulationistic *vs.* non-cumulationistic emergence will be examined in chapter 5.

[34] This issue is examined in chapter 6 in the context of a more detailed investigation of the semantic and ontological problem of self-organization introduced in section 3.3.

		FINITUDE	
		finite	infinite
DYNAMICITY	dynamic	$OE_{[1]}$	$OE_{[2]}$
	static	$CE_{[1]}$	$CE_{[2]}$

Fig 3.4 Summary of environmental kinds.

Using the notation introduced in section 3.4.1 augmented by the temporal parameter *t*, the four kinds of environmental description can be formally defined as follows:

Def^n: open (*OE*) and closed (*CE*) environments (3.4)

$$OE_{[1]} : E_{Cx}(t) \not\equiv E_{Cx}(t+1) \lor E_{Ry}(t) \not\equiv E_{Ry}(t+1) \;\forall\, x, y < \infty$$

$$OE_{[2]} : E_{Cx}(t) \not\equiv E_{Cx}(t+1) \lor E_{Ry}(t) \not\equiv E_{Ry}(t+1) \;\forall\, x, y \leq \infty$$

$$CE_{[1]} : E_{Cx}(t) \equiv E_{Cx}(t+1) \land E_{Ry}(t) \equiv E_{Ry}(t+1) \;\forall\, x, y < \infty$$

$$CE_{[2]} : E_{Cx}(t) \equiv E_{Cx}(t+1) \land E_{Ry}(t) \equiv E_{Ry}(t+1) \;\forall\, x, y \leq \infty$$

Open and closed systems may be defined by analogy with terms in (3.4) as follows:

Def^n: open (*OS*) and closed (*CS*) systems (3.5)

$$OS_{[1]} : S_{Ci}(t) \not\equiv S_{Ci}(t+1) \lor S_{Rj}(t) \not\equiv S_{Rj}(t+1) \;\forall\, i, j < \infty$$

$$OS_{[2]} : S_{Ci}(t) \not\equiv S_{Ci}(t+1) \lor S_{Rj}(t) \not\equiv S_{Rj}(t+1) \;\forall\, i, j \leq \infty$$

$$CS_{[1]} : S_{Ci}(t) \equiv S_{Ci}(t+1) \land S_{Rj}(t) \equiv S_{Rj}(t+1) \;\forall\, i, j < \infty$$

$$CS_{[2]} : S_{Ci}(t) \equiv S_{Ci}(t+1) \land S_{Rj}(t) \equiv S_{Rj}(t+1) \;\forall\, i, j \leq \infty$$

The system-environment coupling relation has two basic forms which describe the extent to which the system may be viewed as partially- or *relatively*-closed:

(1) WEAK coupling (relatively-closed)

$$\{I_{Ra}\ ,\ O_{Rb}\ \forall\ a,b\ |\ a\rightarrow 0,\ b\rightarrow 0\}$$

(2) STRONG coupling (relatively-open)

$$\{I_{Ra}\ ,\ O_{Rb}\ \forall\ a,b\ |\ a\rightarrow \infty \text{ or } b\rightarrow \infty\}$$

There are a number of interesting permutations of the system-environment relation which are worthy of consideration. The most usual is where $E \neq S$, that is, the environment is distinct from the system and the two are connected via an input-output relation I_R-O_R (section 3.4.1). Meehl et al. (1956) define emergence under this scheme in terms of a *movement* within *function space* (that is, the space of all possible system descriptions defined over some finite set of observables). Such movements represent changes in the functional specification of systems and entail a connection to concepts such as extrinsic emergence (section 3.8.1) and emergence-relative-to-a-model (3.6.1). A more interesting relation occurs when $E \equiv S$, that is, when system and environment are identical and either (i) finite or (ii) infinite but algorithmically-compressible[35]. In this case external distinctions between emergent and substrate cannot be made. In terms of the scheme due to Meehl et al., function space itself becomes functionally-specifiable (via a meta-function) and there can be no emergence since there is no movement within function space. The entity can still be regarded as self-organizing if internal distinctions are allowed; this corresponds to a form of intrinsic emergence (section 3.8.1). However, it is unclear how and why internal distinctions arise. Various schemes have been suggested in the literature, for example, symmetry breaking (Jantsch,80) (Silberstein,98) and the calculus of indications (Spencer,69). Nonetheless, the issue is far from resolved since such schemes necessarily make recourse to contestable metaphysical assumptions (for example, tychism or ontical randomness).

The above definitions of closure with respect to systems and environments allow four classes of emergence to be identified (Fig 3.5).

---

[35] This latter possibility raises an important question, viz. is it the environment *itself* that is algorithmically-compressible - in which case "strong" computationalism, viz. reality is computational, must be true - or is it merely the case that *description(s)* of the environment are algorithmically-compressible ?

		SYSTEM	
		open	closed
ENVIRONMENT	closed	I	II
	open	III	IV

Fig 3.5 Classes of emergence.

Examples of class I and II emergence include processes occurring in formal and computational systems (section 3.6). Simon (1981) maintains that closed systems contain the *emergent* (that is, the phenomenal product of emergence) in the substrate, an intrinsically Platonic-Aristotelian view of emergence as the *unfolding* of form (potentiality) in matter (actuality). Examples of emergence in classes III and IV are much more controversial: Formalization of systems with open environments is impossible since *by definition* they are never complete or fully specifiable.

3.9.2. From Emergence to Emergentism

Most current theories of emergence are concerned almost exclusively with epistemological issues as against ontological issues, with the `knowing' of emergence as opposed to the `being' of emergence. For example, Nagel (1961) distinguishes emergence as a thesis about the nonpredictability of certain properties of things, from emergence as the temporal, cosmogonic process described in the work of writers such as Morgan, Alexander and Smuts. Simon (1971) argues that emergence either refers to that which is (at least statistically) predictable in principle or is a non-scientific concept, thus supporting the position held by Ayala (section 3.8.3). Nearly all theories involve the tacit adoption of a set of underlying materialist metaphysical assumptions, viz. determinism, mechanism, reductionism (ontological and methodological if not epistemological). The necessity of a physical (material) substrate for emergence is maintained on the grounds that observed changes in emergents follow observed changes in the material substrate (Kenyon,41). Non-materialistic theories of emergence are rejected on the grounds that "emergents are [thereby] regarded as spiritual creations emanating from an unknown shadowy world." (p.48)

However, alternative conceptions of emergence based on different metaphysical assumptions have been proposed in the literature. For example, Elstob (1986) argues for a theory of emergent evolution which both grounds and is grounded in an indeterministic metaphysics contrasting with the earlier assertions of Pepper (1926) and Nagel (1961) that "neither a belief in indeterminism nor in teleological causation is essential to emergent evolution." (p.377) Elstob (1984) holds that emergent levels are causally-closed and hence, relatively independent of other levels in the systemic hierarchy, essentially a restatement of the position originally articulated by Smuts (1926). The emergence of a phenomenal (systemic) level is viewed as contingently dependent on global context (supersystemic level) and necessarily dependent on components (subsystemic level). However, Elstob maintains that both contextual and component dependencies are relatively weak and of secondary importance when compared with the strong interactions occurring at the phenomenal level, thereby contesting the view that systems are `nearly-decomposable' (Simon,81) (section 3.4.4). Properties (and hence, phenomena) arise from interactions occurring within specific contexts. A stable system exhibiting characteristic properties is produced when a complex of interactions attains causal closure. However, Kim (1992) rejects this view on materialist grounds. On his view, causality is a bottom-up phenomenon which occurs *only* at the most primitive level in the phenomenal hierarchy, viz. the material or physical level[36]. As he states,

> there are no causal powers that magically emerge at a higher-level and of which there is no accounting in terms of lower-level properties and their causal powers and nomic connections (p.18)

The incompatibility of the above two views is a consequence of the incompatability that arises from the adoption of two different metaphysical schemes (ontologies). Elstob's[37] position is interesting because it is consistent with the possibility of `downwards causation' in which systems emerging from a substrate via bottom-up causal processes modify the substrate via top-down causal processes which come into existence at the emergent level. On this view, causality is not a single level phenomenon; rather it is level specific. This is important because it means that epistemological reductionism of higher level phenomena to phenomena at some primitive or substrate level is impossible. (However, as stated previously, there are problems with such non-cumulationist schemes as will be shown in chapter 5.)

3.9.3. Downwards Causation

According to Mayr (1982), the two most important characteristics of emergence are as follows:

1. that the wholes generated can in turn become parts of higher-level systems; thus, emergents are `janus-faced' and can act as *holons*, that is, as both parts and wholes, (Koestler,74).

---

36 In short, the physical universe is held to be causally-closed (Chalmers,96).

37 Elstob's position is critically examined in chapter 6 in connection with a discussion of arguments supporting Heideggerian pluralistic emergentism.

2. wholes can affect the properties of lower-level parts.

Feature (1) is relatively uncontroversial; as an instance of (2), Campbell (1974) cites systems undergoing natural selection: Environmental selection forces operating at the phenotypic level determine, in part, distribution at the genotypic level. Campbell maintains that causation must, therefore, be interpreted from two complementary perspectives, viz. (i) *upwards* - direct, instantaneous (synchronic), physical, and (ii) *downwards* - indirect, historical (diachronic), selective and cybernetic. Klee (1984) has contested this position, maintaining that "there is a difference that is being ignored in this case between determinative connections between levels in a system, and determinative connections between two independently functioning systems .. the nature of the determinative connection between, for example, any one of cells in my body and the organ it is a micro-part of is not the same as the determinative connection, *if any*, between myself and another person [or any other object external to my body]." (p.58) Notwithstanding this criticism, Klee accepts that feature (2) is open to at least three different interpretations described in the literature, viz. downward causation or `retroactive causality' (Elstob,84) via (i) dual control, (ii) global constraints or (iii) supplementary laws, alternative schemes which Davies (1987) has identified with three types of organizing principle, viz. (i) weak, (ii) logical, and (iii) strong.

3.9.3.1. Weak (dual control)

Weak organizing principles are formulated in terms of external constraints, boundary conditions, initial conditions, degree of non-linearity, degree of feedback, distance from equilibrium etc and are statements about the general way in which systems tend to self-organize (Davies,87).

Pattee (1973b) identifies two types of constraint both of which are arbitrary and give rise to a hierarchical order in systems, viz. structural and control constraints, the latter producing control hierarchies via dynamic control of lower level system components. It is important to clarify the distinction between a control hierarchy and the notion of a near-decomposable system hierarchy due to Simon (1981). According to Pattee,

> in a control hierarchy, the upper level exerts a *specific*, dynamic constraint on the details of the motion at the lower level, so that the fast dynamics of the motion at the lower level cannot simply be averaged out. The collection of subunits that forms the upper level in a structural hierarchy now also acts as a constraint on the motions of selected individual subunits. This amounts to a feedback path between levels [emphasis added]. (p.77)

Pattee locates the source of hierarchical control in decision-making or *measurement*, that is, the classification of alternatives according to rules of constraint which are the result of local and arbitrary structures and which act non-holonomically to reduce the number of degrees of freedom in the system (Pattee,72). Classification takes the form of the `writing' of a *record* which, on his view, is a statistical (and hence, irreversible) process (Pattee,71). However, the origin of such records and the constraints which generate them is an unsolved problem. Although Pattee agrees with Simon that control constraints must emerge as a statistical property of interactions between system components, he differs as to the nature of the reverse relation by arguing in favour of downwards causation from the higher to lower level(s). Thus, on Pattee's view, there are at least two conceptions of systems: (i) near-decomposable systems in which higher level behaviour is statistically emergent from the interactions of lower level components, and (ii) constrained systems in which the `interface' between levels organizes the individual dynamics of the lower-level components by a process of classification involving measurement.

Elstob (1984) formulates an interpretation of downwards causation, viz. retroactive causality, in terms of the relative independence of levels in an emergent hierarchy, viz.

> the emergent level, which evolves from the component level but which has a causality that is relatively independent of the component level, can create conditions that give rise to component level interactions that would not occur in the absence of the emergent level. (p.87)

Polanyi (1965) presents a scheme in which the boundary conditions left open by a lower level principle are determined by a higher principle[38], for example, the boundary conditions of the laws of mechanics may be controlled by the operational principles which define a machine; in (Polanyi,66) this is described as the `principle of marginal control'. A similar position is argued by Simon (1971): (Functional) mechanisms are ontically but not epistemically reductive; a machine is consistent (or compatible) with the physico-chemical laws, but not uniquely determined by such laws because it represents only one of a large number of possible arrangements of matter. Thus, it is not possible to deduce the particular structure or/and function of a machine from the physico-chemical principles alone. The nature of the relation between lower and higher level principles is described in (Polanyi,65) as follows:

- The higher principles which characterize a comprehensive entity cannot be defined in terms of the laws that apply to its parts in themselves.

- The operations of higher principles rely quite generally on the action of the laws governing lower levels and are compatible with such laws.

Thus, according to Polanyi, certain kinds of things are subject to *dual control*. In (Polanyi,68), two types of boundary condition are identified which may be described as (i) observation boundary conditions (concerned with that which is *bounded*), and (ii) control boundary conditions (concerned with that which is *bounding*). Type (ii) boundary conditions are those which are primarily involved in dual control, the degree of control exercised depending on the exact nature of the conditions. In (Polanyi,66), the control boundary conditions are identified as emergent:

38 Polanyi (1968) locates the source of the ontology of higher level principles in the concept of *fields*.

> If each higher level is to control the boundary conditions left open by the operations of the next lower level, this implies that these boundary conditions are in fact left open by the operations going on at the lower level. In other words, no level can gain control over its own boundary conditions and hence cannot bring into existence a higher level, the operations of which would consist in controlling these boundary conditions. Thus, the logical structure of the hierarchy implies that a higher level can come into existence only through a process not manifest in the lower level, a process which thus qualifies as an emergence. (p.45)

In the above scheme, bottom-up causation is necessary but *not* sufficient for the emergence of higher levels. Thus, *holism* is distinguished from emergence since only in the latter does irreducibility constitute a necessary condition for the occurrence of the phenomenon.

Bechtel (1986) describes a similar concept in the context of a discussion on teleological explanation and the nature of the mapping between reducing (lower) and reduced (higher) levels in a system description: `Background conditions' denote "particular sets of conditions under which the regularities characterized by the reduced level theory hold" (p.33). He goes further to state that background conditions "constrain the lower level components to do the particular tasks needed for the whole system." (p.33) Adopting a position similar to Elstob, Bechtel maintains that systemic interactions are identifiable at three levels, viz. (i) the subsystemic or component level, (ii) the systemic level at which the system functions as a whole, and (iii) the supersystemic or environmental level.

3.9.3.2. Logical (global constraints)

Global constraints constitute the least controversial interpretation of `downward causation': Global system dynamics are the product of local subsystem interaction dynamics (bottom-up causation). However, the global structure constrains the local dynamics by canalizing interactions, selectively `freezing out' certain interactions (top-down `causation'). The nature of the canalization is best understood in system dynamical terms (chapter 2), viz. trajectories (sequences of global system states), attractors (the end state or sequence of end states of a given trajectory) and basins of attraction (the set of all trajectories that converge on a given attractor). The set of all basins of attraction constitutes the basin of attraction field and its topology is that of a `landscape' of branching transient trees rooted on attractors. (The attractor basins constituting the basin of attraction field are discontinuously connected in discrete deterministic systems). The generation of global macrostructure *from* the deterministic interactions of local microstructures places the concept of emergence associated with logical organizing principles in direct opposition to that associated with the weak organizing principle of Polanyi described in section 3.9.3.1. Emergence via logical or global constraint is shown in Fig 3.6 (overleaf).

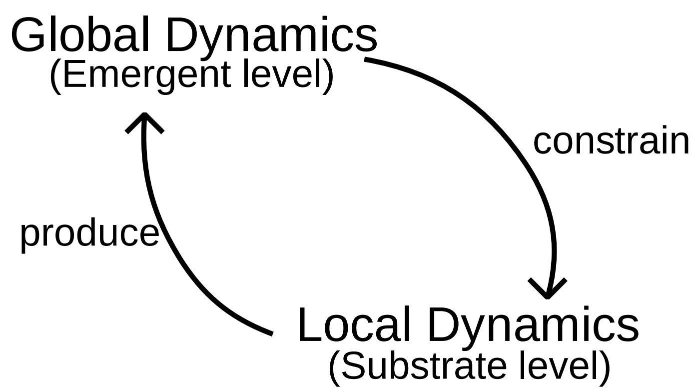

Fig 3.6 Emergence via global constraint.

An alternative formulation of the concept of global constraints is given in terms of `software laws' (Davies,87), logical rules and theorems concerned with system organization, information and complexity independent of specific physical mechanisms; hence, a link between software laws, functionalism and multiple-realizability (chapters 2 and 4). Such laws are not logically deducible from the underlying `hardware' laws traditionally studied in physics although they are compatible with the physical laws. Software laws are organizing principles which `harness' physical laws rather than supplementing them and apply to emergent phenomena, inducing their appearance and controlling their form and behaviour by holistically modifying global system behaviour.

3.9.3.3. Strong (Supplementary Laws)

The concept of supplementary laws may be formulated in two ways: (i) *substitutive* - existing physical laws are modified; (ii) *additive* - existing physical laws are augmented by new laws. This position constitutes the most radical interpretation of downwards causation and necessitates a consideration of the distinction between monistic and pluralistic ontologies. (As stated previously, the issue of cumulationistic *vs.* non-cumulationistic emergence is addressed in detail in chapter 5.)

3.9.4. Pluralism

In the preceding sections, emergent phenomena have been variously described. According to Searle (1992), emergents can be broadly distinguished into two kinds, viz. (1) properties (qualities, features or characteristics) and (2) laws (section 3.7). An alternative classification scheme based on behaviour and briefly described in section 3.5.2 distinguished between structural and functional emergents. However, in the present context, the critical issue is the ontology or being of emergents. There are basically two positions: (1) emergents as apparent and (2) emergents as real[39]. This appearance-reality distinction should not be confused with the epiphenomenality-causality distinction (section 3.9.2), although the two distinctions are related. As stated previously, according to Forrest (1990), apparent-emergents are logically entailed by the bottom-up causation of a computationally-monistic substrate (section 3.6.4). However, the view that emergents are merely epistemological artifacts, that is, appearances, has been contested and two variants of the opposing ontological-pluralist position have been proposed: (i) cumulationistic-pluralism and (ii) non-cumulationistic-pluralism. An early cumulationistic scheme grounded in a space-time event monism was developed by Alexander (1920) and is described in chapter 5 along with a non-cumulationistic scheme due to Bunge (1969). In this section, a recent cumulationistic framework is examined since it is relevant to the debate over pluralism *vs.* monism.

Emmeche et al. (1997) advocate ontological pluralism in the context of a materialistic and evolutionary perspective in which reference is made to "the `local' existence of different ontologies" (p.1). This position is further supported by the assertion that "emergence .. is creation of new properties *regardless* of the substance involved [emphasis added]." (p.5); hence, the view that emergence is "exactly that reasonable aspect of vitalism which [it] is worth[while] to maintain" (p.3). (However, it should be noted that the extent to which ontologies can differ and substances can be disregarded in this scheme is clearly bounded at the outset by the *a priori* commitment to *materialistic* evolution[40].) Emmeche et al. clarify their position as follows:

> if ontologically interpreted .. emergence will characterize the one and only `creative force' in the whole universe and if epistemologically interpreted, it will be a name designating a large scope of various and perhaps very different types of processes. (p.6)

Such processes are hierarchically structured as a series of phenomenal levels in which

> a level is constituted by the interplay between a set of elementary entities and processes acting on a level below (the initiating conditions), constrained by specific boundary conditions (that may have an environmental origin relative to the emerging entities) that determines the `shape' or `form' of the entities at the emerging level. (p.19)

This is consistent with Darwinian (that is, selectionist) accounts of the origin of boundary conditions (Baas,93) (Emmeche,92) in which the environment of a system acts as the observer (section 3.8.1). Furthermore, Emmeche et al. maintain - contrary to Bunge (chapter 5) - that

> levels are *inclusive* .. the psychological level is built upon the biological and the physical, the biological upon the physical: phenomena on one level cannot be reduced to the lower level, but on

---

[39] It should be noted that the reality status of the *process* of emergence is not contested on this position.

[40] Assuming the modern externalistic and mechanistic view of matter described in chapter 4.

> the other hand they can never change the laws of the lower level. Biological phenomena cannot change physical laws - but neither can physical laws as we know them fully explain biological phenomena. The fact that levels are inclusive means that a lower level is a necessary condition for the higher level, and that the higher level supervenes upon the lower. (p.8)

Thus, emergence is viewed as (1) consistent with ontological reductionism and epistemological non-reductionism and (2) a supervenience or coupling relation between phenomenal levels as contrasted with a parallel decoupling of phenomenal levels. However, ontological reductionism does not entail ontological monism since ontological *inclusion* does not preclude ontological *expansion* (section 3.7). With respect to the maxim defining emergence (section 3.2.2), Emmeche et al. maintain that

> what is `more' about the whole is a specific series of spatial and morphological relationships between the parts (p.17)

and these relationships are held to be both epistemologically *and* ontologically significant. As they go on to assert,

> ontologically, there exist other entities than elementary particles. These entities are no less material or materialistic existing than elementary particles. By *materialistic* we only mean that these entities exist independent of a human subject, that is without any subject having thought of, measured or otherwise related itself to the entity. They exist not only epistemologically but also ontologically - understood as independent, objective, and materialistic existence - one might as well say realistic - without reducibility to elementary particles. (p.17)

There are a number of problems associated with this version of ontological pluralism: First, that reality includes the ontologically-subjective as well as the ontologically-objective (Searle,92;95) is a fact which Emmeche et al. accept, yet, it will be argued, do not fully appreciate in the context of the `hard problem' (Chalmers,96) of consciousness; second, and relatedly, this type of pluralistic emergentism necessitates a form of categorial *creatio ex nihilo* (that is, creation from nothing) which is problematic given the adoption of the Lucretian maxim, *ex nihilo nihil fit* (conventionally interpreted as the assertion that nothing comes from nothing) and the First Law of Thermodynamics (chapter 4)[41].

[41] The relation between emergence and the *ex nihilo* maxims will be examined in chapter 6.