Chapter 5

> Life needs something to live on, intelligence needs something to think on, and it is this seething infomation matrix which CAs can provide.
>
> If AI is the surfer, CA is the sea.
>
> Rudy Rucker.

Unification

5.1 Overview

In this chapter, a unifying framework of emergent artificiality grounded in a computationalist metaphysics is presented. First, an approach to unification based on Simon's concept of artificiality and briefly described in chapter 4 is re-examined. A problem with this approach is reviewed and an appropriate solution strategy outlined. The notion of unification is investigated and a brief survey of various approaches to unification described in the literature is presented. It is argued that the requirement for ontological-monism eliminates most schemes from consideration as the basis of a unifying framework of computationally emergent artificiality. However, a metaphysical framework developed by Samuel Alexander (1920) is presented as providing the foundation upon which a computational framework can be constructed; detailed investigation of Alexander's metaphysics and the establishment of an isomorphism between his Space-Time event ontology and CA-computationalism is shown to support this view. Various issues associated with the emergence of artificialities (that is, artificial analogues of natural phenomena) in a CA substrate are examined in the context of a unifying framework of computationally emergent artificiality. Finally, a number of outstanding problems associated with Alexanderian metaphysics are shown to be capable of resolution under CA-computationalism.

5.2. Towards Unifying Artificiality

In chapter 1 it was maintained that if computationalism is to ground artificiality, it must be shown how individual artificialities (that is, artifactual analogues of natural phenomena) can be integrated into a unified framework. The following two arguments were advanced in support of this contention, viz.

1. Computationalism is a metaphysical position and *by definition* (chapter 1) must support phenomenal unification, and

2. Functionalism, which provides the philosophical basis for the possibility of artificiality (chapter 4), necessitates an *isomorphism* or one-to-one correspondence (functional, behavioural, causal, structural or otherwise) between artificiality and naturality to be established in order to grant the former the status afforded the latter, viz. reality (as opposed to mere appearance).

The second reason necessitates consideration of (i) *whether* naturality (or nature) is organized; (ii) if organized, *what* shape or form this organization takes; and finally, (iii) *how* this organization comes into existence. As to the first of these issues, it is crucial to appreciate that the existence of structure, pattern, organization, order etc in nature is a necessary condition for the possibility of science (section 5.3.1) since if nature is to be understood scientifically, it must be assumed (*a priori*) that it *is* organized. Hence, and consistent with the second of the two reasons listed above, viz a commitment to functionalism, artificiality *must* also be organized; moreover, organized so as to be isomorphic (in some sense) with naturality. Regarding the form of this organization, there are a number of possibilities and in the first part of this chapter various schemes will briefly be examined. However, from the outset it should be appreciated that the most widely accepted is that of the *hierarchy* (chapter 3), viz. a system of phenomena *arranged* in a graded order: For example, the naturalistic hierarchy matter→life→mind in which matter is historically antecedent to and the basis of life (and mind) and life is historically antecedent to and the basis of mind. If a hierarchical naturality is valid then a hierarchical artificiality is necessitated as a consequence of the requirement for (some kind of) isomorphism between the two domains and with the organizational `shape' of the former (naturality) determining that of the latter (artificiality).

In chapters 1 and 4, it was stated that functionality can be conceived as originating in essentially one of two ways, viz. as a result of design or as a product of evolution[1]. Similarly, in chapters 3 and 4, it was maintained that there are essentially two ways in which to generate a hierarchy, that is, by design or via emergence. In the context of the current discussion, emergence can briefly be equated with the appearance of a new property (the *emergent*) arising in a complex as a consequence of a specific pattern of activity of its elements (the *substrate*). In order for a property to be considered emergent, it must not be deducible from the properties of the substrate (chapter 3). The two possibilities for hierarchy generation, viz. design or emergence, correspond to the two possibilities for naturality, that is, creation or evolution (chapter 6). Evolution must have occurred if a non-teleological, that is, non-creationistic, explanation for nature is correct since natural phenomena clearly exist and their existence must be accounted for by some

[1] The concept of evolution is briefly examined in chapter 6 in connection with an investigation of the ontical notion of *poiēsis*.

means (chapter 4). (Alternatively, such phenomena could be granted ontologically-primordial status; however, such a position is generally regarded as stultifying within the reductionist Western scientific-philosophical tradition which seeks to explain things in terms of other *smaller*[2] things, a metaphysical commitment which may be traced back to Greek atomism as was shown in chapter 2). Hence, the naturality hierarchy must be emergent and, as a consequence of the necessity for artificiality to be isomorphic with naturality, the artificiality hierarchy must also be emergent. Furthermore, and as stated in chapters 1 and 2, computationalism is an ontology; hence, the artificiality hierarchy must `come forth' or emerge from the computational substrate. Creationism is inconsistent with computationalism since the existence of a creator would negate the ontological (or primordial) status of computation. The *poiēsis* (coming-forth) of computational artificiality is described in this chapter and evaluated in Part II of this study.

5.2.1. Artificiality as Unifier and Unified

In chapter 4, Simon's concept of artificiality was investigated. It was seen that artificiality can be understood in (at least) two senses on his view: First, as denoting the principle by which domains (such as rational thought and evolution by natural selection) might be unified; and second, as a generic term for artificial (as artifactual) analogues of natural phenomena[3]. In this latter sense, artificiality denotes a universal class or category with artificialities as particulars or members of this class. In summary, artificiality as principle (or *unifier*) is responsible for integrating (or *unifying*) artificialities as domains (or the *unified*). The concept of artificiality was shown to be closely linked to a number of key notions associated with computationalism; for example, functionalism (chapter 1) and multiple-realizability (chapter 4). However, Simon's concept of artificiality, and as a corollary, his approach to unification, is problematic when viewed from the perspective of the unification of artificialities. This is because Simon defines artificiality in terms of systemic adaptation to environment and adaptation is a *teleological* notion which does not readily extend `downwards' to matter, the posited lowest phenomenal level in the artificiality hierarchy[4].

---

2 The scientific-philosophical perspective contrasts strongly with the orthodox *theological* perspective in which the ultimate (onto-theological) explanation for things is not the smallest or most atomic but the largest and most encompassing, that is, God. This position holds for most strains of mysticism as well (chapter 1). It will be shown that the atomism associated with mainstream Western science and philosophy also contrasts strongly with Heidegger's ontology of Being and the concept of *incipience*, that is, originary *poiētic* emergence devloped and presented in this thesis (chapter 6).

3 Hence, the `sciences' of the artificial (Simon,69) (Simon,80) (Simon,81).

4 This argument appears to hold even if the teleology of adaptation is teleonomically *a posteriori* or merely functional as opposed to intentional (or teleologically *a priori*) since matter is not ordinarily conceived in functionalistic terms (chapter 4); in fact, according to Dennett (1995), functionalism is ultimately *grounded* in materialism (or physicalism).

5.2.2. The Teleology Problem Revisited

An attempt was made in chapter 4 to define artificiality in non-teleological terms so as to resolve the conflict between the concept and a non-teleological view of matter. (Matter must be included within the artificiality hierarchy since it constitutes a basic *phenomenal* kind.) However, as discussed in the section on artificial physics, there have been moves toward the redefinition of matter in computational terms, viz. the computational theory of matter (CTMa) proposed by Wheeler, Fredkin, Toffoli and others. If this approach is valid and if computation is, in some sense, functional (teleological), then matter itself becomes teleological and Simon's artificiality principle *can* be extended throughout the phenomenal hierachy, both upwards *and* downwards. However, there are a number of problems with the artificiality principle as conceived by Simon: For example, even if matter *is* teleological, its teleology is not of the same order (or degree) as that associated with the teleological primitives of his scheme (which are either mental-rational or biological); it is simply the case that Simon's primitives are not sufficiently primitive to realize the teleology of matter[5], a teleology which if anything (according to current systems-theoretical thinking) takes the form of self-organization (chapters 3 and 6). Alternative schemes have been proposed which attempt to address this problem. For example, Campbell (1985) presents a framework (Fig 5.1) in which the phenomenon of evolution is variously realized across a broad teleological spectrum ranging from simple self-organization (physical systems) through natural selection (*a posteriori* or *as-if* intentionality associated with biological systems) and culminating in self-adaptation (*a priori* or genuine intentionality associated with conscious, volitional entities). According to Campbell, as the evolutionary hierarchy is ascended, new evolutionary processes evolve which support increased autonomy and new kinds of teleology; hence, a distinction is made between intra-phenomenal evolution and inter-phenomenal or meta-evolution in which new teleological mechanisms arise.

5.3. Towards A Unification of Artificiality

Although schemes such as those due to Campbell and others address many of the problems associated with the conventional approach to unifying artificialities, various problems remain outstanding. Perhaps the most immediate is the need to clarify what is entailed by *unification*. Consequently, in this section the notion of unification is briefly examined. (It should be noted at the outset that a more detailed review of the literature on unification is beyond both the aim and scope of this study.) Various scientific and philosophical works have dealt with the issue explicitly; however, it is implicit in any attempt at explaining the phenomenal world.

---

5 That is, not unless one is willing to concede the `dumbing down' of living matter to non-living matter, involving what might be described as a form of `property-hiding'. However, this position is unacceptable since it violates the very essence of conventional emergentism which is loosely captured in the maxim "more from less".

increasing self-referentiality ↑

		Organization	Emergent Property	Analysis Format
AI	self-adaptation	Information about future self	Future causality	Future self-reference
AL	natural selection	Information about self Information	recursive causes are their effects	Self reference Cybernetics
AP	self-organization	Entropy/Negentropy Mechanical objects None (elementary particles only)	Unidirectional cause and effect Deterministic cause and effect Acausal	Thermodynamics Newtonian mechanics Quantum mechanics

Fig 5.1 Campbell's Evolutionary Hierarchy.

5.3.1. Unification as Idea and Ideal

Unification as an *idea* is related to the metaphysical view that the world[6] constitutes (in some way) a connected totality, a whole or unity and its origins in the Western tradition are traceable at least to the Ionian Greeks (Collingwood,45). The belief in the unity of nature is a cornerstone of modern science, specifically of physics, and, as Barrow (1991) states, is aesthetically motivated by a heuristic maxim or *ideal*, viz. Ockham's Razor, which normatively identifies the simplest *possible* explanation of a thing with the *actual* explanation of the thing. Interestingly, it is an appeal to Ockham's Razor and other aesthetic criteria such as beauty and truth - what Barrow refers to as `prospective elements' - which leads him to argue against a computational or mathematical theory theory of everything. Appealing to the Gödel incompleteness results within logic and the existence or Platonic subsistence of uncomputable numbers (chapter 2) in support of this position, Barrow maintains that "no non-poetic account of reality can be complete". As he states,

> unlike many others that we can imagine, our world contains prospective elements [such as beauty, simplicity and truth]. Theories of everything can make no impression upon predicting these prospective attributes of reality; yet, strangely, many of these qualities will themselves be employed in the human selection and approval of an aesthetically acceptable Theory of Everything. There is no formula that can deliver all truth, all harmony, all simplicity. No Theory of Everything can ever

6 In fact, use of the term `the world' and its identification with "the totality of all that exists" conceals a tacit *a priori* commitment to metaphysical unification in the assumption that such a totality is both meaningful and real (that is, existent). For example, if a *processualist* (chapter 2) interpretation of existence (or Being) is adopted then the *substantialist* notion of `the world' as the totality of existents *at an instant* is undermined: The interpretation of `world' as identical to instantaneous universal state commits, according to this latter view, what Whitehead (1926) has called the Fallacy of Misplaced Concreteness, that is, "the accidental error of mistaking the abstract for the concrete." (p.66) I am grateful to Mike Elstob for the above point.

> provide total insight. For, to see through everything, would leave us seeing nothing at all. (p.210)

Newell's (1990) views on the nature of unification, expressed in the context of a consideration of unified theories of cognition, may be generalized and extended to other phenomenal kinds. As he states,

> a unified theory will unify our existing understanding of [a phenomenon]. It will not be a brand-new theory that replaces current work at every turn. Rather, it will put together and synthesize what we know. On the other hand, it can't just be a pastiche, in which disparate formulations are strung together with some sort of conceptual bailing wire. Its parts must work together. (p.16)

Newell lists the following as characteristic of unification, viz.

1. Unification is always an aim of science
2. Bringing all parts to bear on a behaviour
3. Bringing multiple constraints to bear
4. Increased rate of cumulation
5. Increased identifiability
6. Amortization of theoretical constructs
7. Open the way to applications
8. Change from discriminative to approximative style
9. Solve the irrelevant-specification problem (p.18)

Feature (6) is of particular interest since it is consistent with the logical end result of repeated application of Ockham's Razor, viz. *monism*, of which computationalism is a particular kind.

5.3.2. Issues in Unification

In this section, various approaches to unification are examined. The work of systems theorist Mario Bunge, who has investigated concepts such as emergence, levels and hierarchies in detail, provides a suitable interpretative framework within which to examine the different unification strategies. His characterization of each approach is simple, systematic and rigorous and will be followed in this study. However, before examining Bunge's scheme, it is worthwhile briefly considering some other views in order to gain an appreciation of the various positions within the movement towards scientific or/and philosophical unification. For example, Polanyi (1966) argues for "a picture of the universe filled with a strata of realities, joined together meaningfully in pairs of higher and lower strata." (p.35) Clearly, this statement entails support for a *hierarchical* view of nature (and by extension, in the context of this study, of artificiality). Gerard (1969) maintains that *systems* (chapter 3) both imply and are implied by hierarchies; consequently, nature (and artificiality) *must* be hierarchically structured if nature consitutes a unity, that is, system. This follows from the fact that

> `system' implies an entity containing subordinate units in some relationship to each other; and that implies hierarchy, superordinate and subordinate levels. Conversely, hierarchy implies units related to each other at different levels; and that is a system. (p.226)

Pattee (1969), who is interested in *autonomous* hierarchy production, that is, the emergence of hierarchies in closed systems (chapter 3), lists three conditions for a hierarchy, viz.

1. autonomy, that is, a closed physical system
2. elements in the system which obey laws of physics
3. collections of elements which constrain individual elements (p.163)

Significantly, condition (2) is consistent with ontological reduction (although, not necessarily to a monism). Rosen (1969) offers the following interpretation of hierarchies addressing both the epistemological and ontological aspects of the concept:

> A hierarchically organized system is simply one which is (a) engaged simultaneously in a variety of distinguishable activities, for which we wish to account, and (b) such that different kinds of system specification or description are appropriate to the study of these several activities. (pp.179-180)

This position is consistent with Rosen's and Cariani's emergence-relative-to-a-model concept (chapter 3) and as Rosen states, "the idea of a hierarchical organization simply does not arise if the same kind of system description is appropriate for all [functions or behaviours]." (p.180) He lists three requirements for solving the problem of hierarchical organization in the context of a discussion of thermodynamics:

1. The *universality* of the underlying microdynamics, viz. any aspect of system behaviour can, in principle, be expressed in terms of the micro-description.

2. A determination of how the state variables of the macro-description could actually be described in terms of the microdynamics, i.e. in terms of the observables of the microsystem. This is a non-trival requirement and is a necessary condition for universality.

3. The implementation of (2) to actually derive the kinetic properties of the macrosystem from those of the microsystem.

In this example (thermodynamics), (2) and (3) are what statistical mechanics does. Rosen is, however, sceptical about the success of applying the above in the context of life and mind on the grounds that a very specific *historical* ordering of concepts made possible the above reduction: First, a phenomenological specification of *macrosystem* behaviour (the gas laws); second, a specification of the *microsystem* dynamics (Newtonian mechanics); and third, a *connecting* formalism (statistical mechanics). As he states,

> if the gas laws had not been known *first*, they would never have been discovered through statistical mechanics alone. Formalism will indeed enable you to form any averages you want, but it will not tell you what these averages mean, and which of them are useful and important in specifying and describing macrosystem behaviour. (p.187)

This position is supported by Gerard (1969), viz. "it is perfectly meaningless to measure something, with higher and higher degrees of precision, if the thing you measure is more or less meaningless." (p.219) Thus, for Rosen (and Cariani), hierarchies are *epistemological* constructs[7], a view which is supported by Tallis (1994) who argues against observer-independent hierarchical-realism in the context of a discussion of the mind-body problem.

Berlinkski (1986) argues for a stronger view, maintaining that biology is not reducible to physics and citing Darwinian evolutionary theory in support of his position:

> The usual Darwinian concepts of fitness and selection [are concepts] that do not figure in standard accounts of biochemistry, which very sensibly treat of valences and bonding angles, enzymes and metabolic pathways, fats and polymers - anything but fitness and natural selection. (p.235)

Following Rosen, he further states that "the standard and, indeed, the *sole example of reduction successfully achieved* involves the derivation of thermodynamics from statistical mechanics." (p.234) Consistent with this position, Miller (1995) is led to criticize what he views as "messianic predictions that theoretical biology will be revolutionized, perhaps with emergence replacing evolution as the central explanatory principle of life." (p.16) However, Kauffman (1993, 1995) has contested the exclusiveness of natural selection in evolution maintaining that self-organization may in fact play the dominant role in generating the phylogenetic hierarchy; thus, the stage is set, at least *in principle*, for a reduction of neo-Darwinian theory to some simpler theoretical framework. Finally, mention might be made of Laszlo (1993) who maintains that a unified conception of the world is scientifically possible based on "a concept of reality organically shaped by interacting universal fields" (p.26). According to this view,

> science .. may have already reached the portals of real insight into the unitary interactive process that generates the diverse and consistent orders of the distinct yet not categorically discrete realms we customarily identify as `matter', `life', and `mind'. (p.20)

5.3.3. Bunge's Approaches to Unification

The underlying notion which unifies all the above positions is that of hierarchy, irrespective of whether hierarchies are viewed as ontological or epistemological, real or apparent. However, other schemes for phenomenal unification have been proposed in the literature and it is important to have some appreciation of their character.

According to Bunge (1963), if the notion of a phenomenal *level* is accepted as phenomenologically (but not necessarily epistemologically or ontologically) primitive, nine interpretations of the level concept can be distinguished: (pp.36-48)

1. *Degree* (qualitative difference is not a necessary condition)
   Def^{n}: An object belongs to a degree D_n higher than another degree D_{n-1} if and only if it surpasses in at least one respect all the objects belonging to the lower degree.

---

7 As will be seen in chapter 7, this position is supported by Maturana (1997).

(1) $D_n=D'_{n-1}$; (2) $D_n>D_{n-1}$

2. *Degree of Complexity* (qualitative difference is not a necessary condition)

Def^n: An object belongs to a degree of complexity C_n higher than another C_{n-1} if and only if the number of its constituents and consequently the number of its interrelations is larger than both the number of elements and mutual relations of the objects belonging to the lower degree.

(1-2) as above; (3) $|x|(C_n x)>|x|(C_{n-1}x)$

3. *Degree of Analytic Depth*

Def^n: A piece of knowledge (description, hypothesis, theory, technique, method) belongs to a degree of analysis A_n deeper than another A_{n-1} if and if it accounts for a larger number of features of the referents common to both pieces of knowledge or if it explains some properties occurring in A_{n-1} in terms of concepts peculiar to A_n or if it decomposes its objects more thoroughly than A_{n-1} does or if it reveals a finer mesh of relations.

(1-2) as above; (4) $A_{n-1}(P\cup S)A_n$, where P denotes \`is part of' and $S$ denotes \`is subsumed under'

4. *Emergent Whole*

Def^n: An emergent whole is an entity that, in some respects, behaves as a unit; if complex, it is highly integrated and has qualities which its parts lack; and it arises from lower order units and may give rise to higher order emergent wholes.

(1-3) as above; $W_n(P\cap S)W_{n+1}$, where P denotes \`is part of' and $S$ denotes \`is subsumed under'

5. *Poistem*

Def^n: A poistem is a system of interrelated qualities or variables. Symbolically, the n-th poistem is the n-th set of qualities $P_n = \{Q_i\}_n$.

(6) $(i)(\exists n)Q_i\in P_n$; (7) $(\exists i)(\exists k)(\exists R)Q_iRQ_k$; (8) $P_n\cdot P_{n+1}\neq 0$

6. *Rank* (top-down graded hierarchy)

Def^n: A rank (or grade in a hierarchy) is an element in a discrete linear sequence such that its status is higher or lower than the neighbouring ranks and such that, unless it is the highest of all, it is dependent in some respects on the higher ranks.

(1) as above; (9) $R_n>R_{n+1}$; (10) $R_m\cdot R_n=0,\ m\neq n$; (11) $R_n\ dep\ R_{n-1},\ n\neq 1$, where \`*dep*' denotes dependency; (12) $|x|(R_{n-1}x)<|x|(R_n x)$

7. *Layer*

Def^n: A layer or stratum is a section of reality characterized by emergent qualities. Symbolically: $S_n=\{Q_n\}$, where \`Q_n' designates one of the *nova* peculiar to S_n.

(1) as above; (13) S_{n-1} prec S_n, where \`*prec*' denotes a precedence relation (temporal, causal, logical etc); (14) $S_m\cdot S_n=0,\ m\neq n$

8. *Rooted Layer* (cumulative, superpositional)

Def^n: An object belongs to a rooted layer Y_n higher than another Y_{n-1} if and only if, in addition to all the qualities that characterize Y_{n-1}, it has a set of emergent qualities Q_n of its own.

(1) as above; (15) $Y_n\ em\ Y_{n-1}$, where \`*em*' denotes \`emerges from'; (16) $Y_n=Y_{n-1}+\{Q_n\}$; (17) $Y_m\cdot Y_n\neq 0$

9. *Level (proper)*

Def^n: A level is a section of reality characterized by a set of interlocked properties and laws, some of which are thought to be peculiar to the given domain and to have emerged in time from other (lower or higher) levels existing previously.

(18) L_n *em* L_{n-1} where `*em*' denotes `emerges from'

A graphical representation (Bunge,60) of each concept is shown in Fig 5.2:

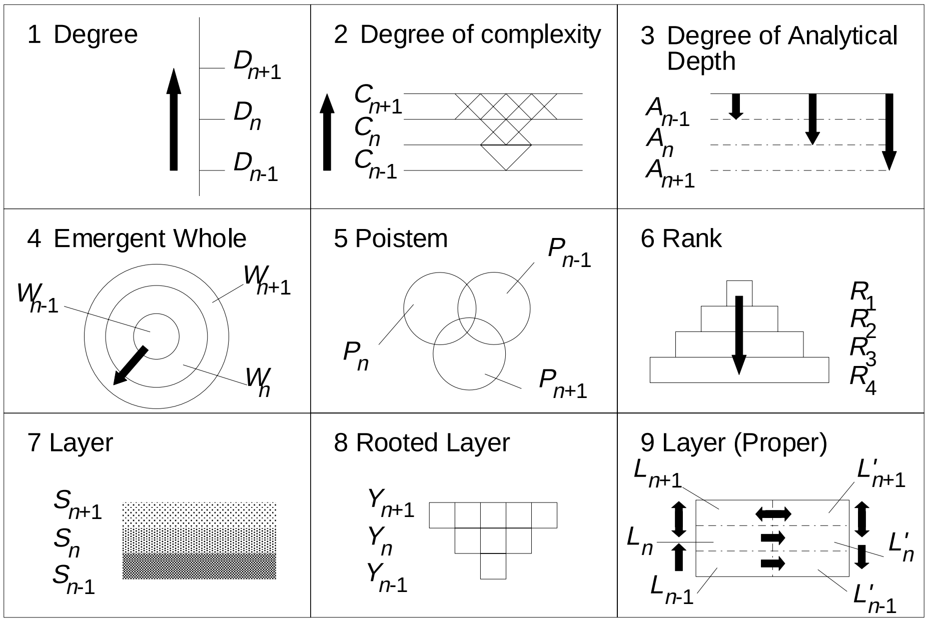


Fig 5.2 Bunge's nine interpretations of the level concept.

Bunge objects to the layer concept (7) maintaining that is unlikely that the new qualities at a level *n* subsist in a way which is completely independent of qualities at *n*-1. However, he also contests the *cumulationist* position which postulates the existence of conservative (or property-preserving) emergent phenomena in the external world, viz. "it would seem that the superposition of patterns occurs in the highest levels only, whereas the spontaneous emergence of qualities in nature is not cumulative, some qualities being lost in the process of emergence." (p.45) This leads Bunge to propose the concept of an emergent layer or organization[8] (9), a complex fusion of the positive aspects of each of the other concepts, most notably that of the emergent whole (4) and the poistem (5). These two concepts are important because they allow for an ontologically non-reductive emergentism[9], which is the position held by Bunge (1969) as shown in the following summary of his metaphysics:

8 Bunge (1969) distinguishes phenomena structured as a series of organizational levels from phenomena structured hierarchically; on his view, hierarchies involve valuative relations of domination and superiority which are absent from organizational levels. This position is consistent with Simon's (1981) distinction between hierarchies and partitionings (chapter 3).

9 As stated in chapter 3, Bunge's emergentism is fundamentally non-cumulationistic.

Ontology (Integrated Pluralist)	*Epistemology* (Realist)	*Methodology*
O_1 Reality (= the world) is a level structure such that every existent belongs to at least one level of that structure.	E_1 The real level structure is knowable and scientific knowledge is a level structure that matches the former.	M_1 Start by limiting your inquiry to one level. Should this level prove insufficient, scratch its surface in search for further levels.
O_2 In the course of every emergence process (self-assembly or evolution) some properties, hence also some laws, are gained while others are lost.	E_2 Every newly formed science has its peculiar objects and special methods. And, although every science retains some of the ideas typical of the parent science(s), it does not preserve them all and it introduces new concepts absent from the latter.	M_2 Face emergence and try to explain it: begin by attempting to explain novelty away but, should this move fail, take it seriously.
O_3 The newer levels depend on the older ones both for their emergence and for their continued existence.	E_3 The understanding of any level is greatly deepened by research into the adjacent levels, particularly the underlying ones.	M_3 Explain the emergence of every level in terms of some of the older levels without skipping any intermediate level.
O_4 Every level has, within bounds, some autonomy and stability.	E_4 Every level of science has, within bounds, some autonomy and stability.	M_4 Begin by investigating your class of facts on their own level(s): introduce further levels only as required.
O_5 Every event is primarily determined in accordance with the set of specific laws that characterize its own level(s) and the contigous levels.	E_5 Every system and every event can be accounted for (described, explained or predicted, as the case may be) primarily in terms of its own levels and the adjoining levels, without necessarily involving the whole level structure.	M_5 Start by finding or applying the intralevel laws. Should this strategy fail, resort to hypothesizing or applying interlevel laws.

Table 5.1 Bunge's Metaphysics

There are (at least) three points to note in connection with the above scheme: First, O_2 is consistent with an *intersectional* (or poistem-like) view of the relations between phenomenal levels. (It is interesting to consider how level intersection might occur; in this respect, the work of Van Gigch (1990) who describes an approach to domain unification based on inter-paradigmatic conflict resolution via epistemological abstraction or *metamodeling*, is particularly relevant.) However, it conflicts with the cumulative *hierarchical* view of the relations between phenomenal levels which is characteristic of other more conventional emergentist frameworks, for example, that due to Alexander (section 5.4); second, O_3 implies strict causal dependency of higher phenomenal levels on lower levels. It could, therefore, be argued that higher phenomenal levels are, in some sense, non-causal or *epiphenomenal* (chapter 3). This issue will be addressed further in connection with an examination of Alexander's metaphysics; third, consistent with conventional views as to the relations between the various forms of reductionism (chapter 3), Wimsatt (1972) maintains that a commitment to epistemological reductionism is entailed in the bias towards theoretical (ontological) monism. However, Bunge (1963) is critical of attempts at epistemological reduction of phenomena to ontologically-monistic substrates - what he refers to as "philosophical Dadaism". He contests this position, arguing that it is possible - in fact, *necessary* - to adopt ontological pluralism while maintaining a commitment to epistemological reductionism: On his view, epistemological reductionism entails *partial* theoretical reduction of a higher level phenomenon to a lower level phenomenon; not only is totalistic reduction not a necessary condition for epistemological reduction, but, according to Bunge, this is simply not possible since ontological pluralism is both phenomenologically and metaphysically correct. Hence, O_5 is problematic because it reinforces Bunge's *a priori* commitment to ontological pluralism, a view which seems to contradict the assertion that his scheme is `scientific' given aesthetic consensus in favour of a link between science and atomistic reductionism as articulated in Ockham's Razor and manifested in the goal of an ontically-monistic theoretical framework. It could be argued that only in a naive or `folk'-phenomenological sense are phenomena self-determining or self-causing; ultimately, the causality associated with all higher level emergent phenomena reduces to the causality of the ontological substrate. (This position is characteristic of monistic ontologies in general and Alexander's framework in particular as will be seen in section 5.4). Although it is tempting to reason that if O_i, then E_i and if O_i and E_i, then M_i, Bunge (1969) offers the following criticism of this line of argument, viz.

> even assuming that our methodology is correct .. we would not be justified in inferring that it verifies the metaphysics and the epistemology behind that methodology .. All we can do is to draw the following *weak* (nondeductive) inferences:
>
> If O_i, then E_i
> Now, E_i
> Hence, maybe O_i
>
> and
>
> If O_i and E_i, then M_i
> Now, M_i
> Hence, maybe O_i and E_i (p.27)

Bunge further maintains that "to the extent to which our methodology works, the ontology and the epistemology behind it look *plausible* .. whether or not our metaphysics and epistemology of levels are actually true, they seem to have been *fruitful*." (p.28) The problem with this view is that M_i *because* E_i; that is, the methodology works *because* epistemological realism has been adopted *a priori*. Hence, there is a circularity involved in the abductive justification of Bunge's metaphysics.

5.3.4. Problems with Bunge's Framework

Bunge's (1969) commitment to ontological-pluralism leads him to postulate the following two metaphysical theses, viz.

1. In the course of every emergence process (self-assembly or evolution) some properties, hence, also some laws, are gained while others are lost.

2. The newer levels depend on the older ones both for their emergence and for their continued existence.

Although the properties lost and gained are those associated with a new phenomenal level, this view is problematic since the whole discourse on properties presupposes a commitment to a substance-predicate metaphysics in which substance remains unchanged regardless of processes undergone; in short, emergent phenomena correspond to the appearance of new properties associated with an *unchanging* noumenal substance or substrate. On this latter view, the vital is *also* physical. If, however, life is an emergent phenomenon which results in the *displacement* of certain physical properties, it may be that the *properties* associated with the emergent are such as to prevent the occurrence of the (physical) *process* which brought the emergent into existence. There are basically two solutions to this problem: (1) embrace a `strong' (or totalistic) decoupling between phenomenal levels or (ii) adopt cumulationism[10]. The former entails a kind of parallelism (as in Spinoza) and is problematic since the connection between emergent levels is unexplainable in even methodologically-reductive terms; the latter leads to a position close to that of Alexander (1920) who is, in fact, able to incorporate Bunge's non-cumulationist arguments into his framework (section 5.4.4).

Yet another problem with Bunge's approach in the context of the unification of artificialities arises as a consequence of his *a priori* commitment to ontological pluralism. As stated in chapters 1 and 2, computationalism is a monistic metaphysics; hence, Bunge's framework is unsuitable as a foundation for unifying *computationally* emergent artificiality (chapter 5). For this reason it is necessary to consider other approaches to unification grounded in a monistic ontology and supporting phenomenal emergence. In this connection, a metaphysical scheme developed by Samuel Alexander and based on a Space-Time event monism is presented as a suitable foundation upon which to construct a unified framework of computationally emergent artificiality. Adoption of this framework is motivated by two considerations: First, according to Alexander (section 5.4), Space-Time is ontological; empirical existents (emergents) are phenomenal. This is consistent with the Kantian appearance-reality or phenomenon-

[10] This position is implicitly endorsed by Maturana and Varela (1980) in their conception of the autopoietic organization as embedded in (and supervenient on) physical space (chapter 6). It is crucial to appreciate that Bunge's non-cumulationism is not *necessarily* self-defeating under a substance-property metaphysics. This is readily shown by distinguishing *necessary* from *contingent* properties; the former *are* cumulationistic (and preserved in the emergent) while the latter *can be* non-cumulationistic (and displaced in the emergent).

noumenon distinction (chapter 1), a distinction which is intrinsic to computationalism (chapter 2) and artificiality (chapter 4); second, it is conceivable that Alexanderian Space-Time is isomorphic with CA-computationalism, that is with a computationalism implemented (or *realized*) by a cellular automaton (chapter 2).

5.4. Alexander's Metaphysics

In this section, the Space-Time ontology of Samuel Alexander is briefly investigated. An examination of this metaphysical scheme is necessary since it describes one of the earliest attempts at unifying phenomena under an emergentist framework in which the notion of pattern or organization (form, structure) is categorially primitive; hence, it establishes a precedent for computationally emergent ontologies such as those based on cellular automata (section 5.5). The following account is drawn from Alexander's *Space, Time and Deity* (1920) and discussions of this work appearing in (Collingwood,45) and (Brettschneider,64).

5.4.1. Introduction

Alexander presents a unified metaphysical framework based on a continuous and infinite Space-Time event ontology. Before presenting a detailed description of this framework, three important aspects will be briefly examined: First, according to Brettschneider (1964), Alexander's metaphysics is *internalist* and *idealist* nothwithstanding claims by Alexander to the contrary, viz. that it is *externalist* (realist) and *empiricist*. Brettschneider justifies this interpretation of Alexanderian metaphysics on the basis of Alexander's commitment to a ontological *coherence* theory of truth (chapter 3). This point is extremely important because it immediately establishes a connection between Alexanderian metaphysics and CA-computationalism. the latter of which is also held[11] to be grounded in an internalist or `intrinsic' (Crutchfield,94) view of truth and reality (chapters 2 and 3).

Secondly, Alexander, like the materialist philosophers before him, makes a distinction between categories and qualities: *categories* are the primordial, that is ontological, properties of space-time; *qualities* associated with empirical existents (or space-time complexes) are phenomenal and have merely (*contingent*) factual status. Moreover, since the categories are ontological, they are non-emergent (that is, *necessary*), and present in all orders of empirical existence whereas qualities are emergent. But *emergent* in what sense ? According to Alexander, quality is pattern- or organization-dependent. Furthermore, qualities are held to be objective: As Collingwood (1945) states, quality

---

11 It is important to distinguish between *topological*-internalism, which *is* being asserted in this context, and *experiential*-internalism, which is not. In chapters 6 and 7, it is argued that computationalism is incapable of supporting the latter as a consequence of the fact that it is a variant of mechanism and hence, ontologically-objective (or externalistically non-experiential).

is "not a mere phenomenon, it does not exist merely because it appears to a mind; it exists as a function of structure in the objective world." (p.160) Alexander goes on to state that "quality belongs to things as mind or consciousness belongs to life-processes of a certain configuration" (Vol.II, p.47) and according to Brettschneider,

> quality is the mind of structure or form .. Just as mind emerges from body, so quality emerges from structure [and] *coherence* is the agent that pulls together the co-presences of space-times that make up qualities .. In Aristotelian terms, the finite patterns of pure Space-Time are material causes; the restlessness of Time is the efficient cause; coherence as the principle or organization of Space-Time is the formal cause; and quality as the emergent unity of Space-Time organization is the final cause [emphasis added]. (p.73)

However, although Alexander unequivocally maintains that emergents are *not* epiphenomenal (or non-causal), the opposite is clearly the case since on his scheme causation is categorial; hence, all causation is bottom-up and all orders of empirical existence are ontically, although not epistemically, reducible to space-time complexes.

Thirdly, Alexander is emphatic in asserting that "quality is not a category but an empirical generalization of the various specific qualities of things, or a collective name for them all" and that "experience does not acquaint us with quality as such; as it does make us acquainted with quantity or substance as such." (Vol.I, pp.326-327) He holds that "quality is to specific qualities as colour is to red, green, and blue" and that "even if [it] could be maintained [that there is a plan of colour], it cannot be held that there is any plan underlying red and hard and life which is modified into these specific qualities." (Vol.I, p.327) Furthermore, he maintains that "complexity in Space-Time makes everything a complex, but not a quality. It is specific sorts of complexes which are hard or sweet. Complexity as such is not a qualitative but a quantitative or purely spatio-temporal determination .. Quality is therefore not categorial but empirical." (Vol.I, pp.327-328) However, as will be argued in chapters 6 and 7, this view is problematic on (at least) two counts: First, quality *can* be universally defined following Nagel's (1979) formulation, viz. quality as internal subjective-experience or *what-it-is-like-ness*; and secondly, Whiteheadian panexperientialism (chapter 1) provides a framework within which specific qualities can be *constructed* according to plans. According to this latter scheme, while it is correct to hold that not *all* complexes are quality-bearing[12], quality is, nonetheless, ontically primitive and hence, both empirical and categorial. Additionally, this position is consistent with Alexander's claim (which, according to Brettschneider, is problematic on his metaphysics) that both the ontological categories and empirical existents can be *experienced*.

---

[12] According to Whitehead (1978) and Griffin (1988, 1998), complexes of actual occasions can assume one of two forms, viz. genuine individuals which are experiential and mere aggregates which are not. However, the ontological primitives in both are experiential, that is, quality-bearing. Consequently, quality is an ontological category.

5.4.2. Space-Time

As stated above, Alexander's metaphysics is based on an infinite and continuous Space-Time event ontology. Before discussing this ontology in detail, it is worthwhile briefly comparing the various views on the nature of space and time. In this respect, a summary of the three main positions due to Bunge (1977b) is particularly informative (Table 5.2):

	Container View	*Prime-stuff View*	*Relational View*
Proponents	naive or commonsense view, Democritus, Newton, Laplace	Clifford, Alexander, Einstein, Wheeler	Aristotle, Leibniz, Mach, Whitehead
Characteristics	spacetime-matter dualism, static space and time, non-physical spacetime, self-existing (absolute)	geometric spacetime monism, self-existing (absolute), things = spacetime	spacetime = network of relations among factual items, viz. things and their changes

Table 5.2 Three views of spacetime

According to Bunge, on the prime-stuff view of spacetime, "there is no need of semantic hypotheses (`correspondence rules') [since] the border line between *formal* science and *factual* science disappears - perhaps also that between constructs and things [emphasis added]." (p.280) This point is critical since it is consistent with the *coherence-theoretical* view of reality implicit in the computationalist position.

For Alexander, reality *is* Space-time or motion itself; a single, self-contained infinite *stuff* taking the form of a continuous plurality of point-instants or *events* which are the ultimate constituents of all things. On this view, empirical things or substances are parts or modes of this stuff. The latter is not to be confused with *substance*; stuff is prior to substance, which is identifiable with physical or empirical existence[13]. Space and Time are totalities which denote the two abstract aspects of the one absolute infinite stuff of Space-Time; space and time denote finite `regions' of Space and Time respectively. According to Alexander, "Time is repeated in Space" and "Space is repeated in Time" (Vol.I, p.48). These one-many relations can be visualized graphically as shown in Fig 5.3:

[13] This point is important on two counts: First, it necessitates a non-substantialist interpretation of Alexander's metaphysics. Hence, Rescher (1996) is apparently justified in identifying Alexander's scheme as processualist (chapter 2); second, and in the context of this thesis, it establishes a precedent for considering ontology (that is, the study of Being) in non-substantialist terms. A scheme based on the Heideggerian notion of Being, a non-substantialist *and* non-processualist ontology (Dreyfus,92), is outlined in chapter 6.

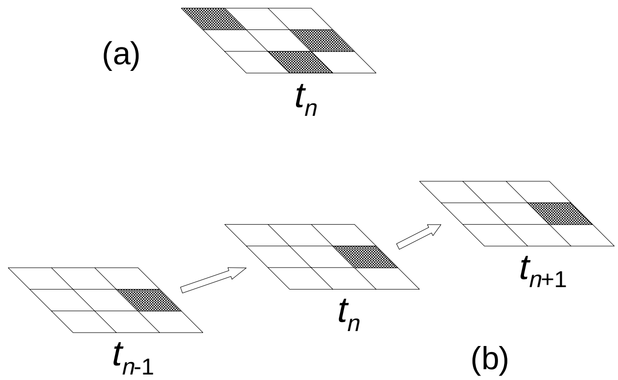


Fig 5.3 (a) Time repeated in Space; (b) Space repeated in Time.

In (a) "Time is repeated in Space", viz. a one-many correspondence between an instant (t_n) and the points it *occupies* ($\{s_i,..,s_j\} \subset S$, where S is the infinite set of all points) is established; in (b) "Space is repeated in Time", viz. a one-many correspondence between a point (s_i) and its *occurrences* over a number of instants ($t_{n-1},..,t_{n+1}$) is established. This framework enables Alexander to formulate a concept of motion. On his scheme,

> points do not of course move in the system of points, but they change their time coefficient. What we ordinarily call motion of a body is the occupation by that body of points which successively become present, so that at each stage the points traversed have different time-values when the line of motion is taken as a whole. (Vol.I, p.61)

The above point can be restated as follows, viz. "the meaning of motion is .. not that the point of space itself moves as if it were a material body shifting its place, but that the time of a point ceases to be present, and the present is transferred to another point continuous with it." (Vol.I, p.272) As Brettscheider (1964) observes, "a stretch or block of space moves along as it changes its time-coefficients." (p.24) Alexander describes "a grouping or complex of point-instants or pure events [as] a configuration of space-time or of motion." (Vol.I, p.210) This conception of motion provides the basis for his view of objects, that is, empirical existents: According to Alexander, finite ontological existents are

> continuously connected groupings of motions, and they are connected through the circumambient Space-Time with other such groupings or complexes. In less metaphorical language, they are complexes of motion differentiated within the one all-containing and all-encompassing system of motion [that is, the system of Space-Time]. (Vol.I, p.138)

According to Brettschneider (1964), "an object or entity is, in his terms, a sequence of Space-Time relations." (p.27) Furthermore, "an object should not be thought of as a static, unchanging entity despite appearances to the contrary. Instead it should be conceived of as a process." (p.25) Thus, Alexanderian ontology is fundamentally *processual* (chapter 2) in character. Alexander goes on to assert that

> in any point-instant the instant is the mind or soul of its point; in a group of points there is a mind of those points, which upon the primary level of Space-Time itself is the corresponding time of that complex. (Vol.II, p.39)

Hence, the statement that `Time is the mind of Space and Space is the body of Time'. It should not, however, be understood from these statements that mind is primordial; Alexander uses these expressions in a purely analogical or *functionalistic* (chapters 1 and 4) sense. As he states,

> I do not mean as Leibniz meant that things on their different levels possess varying degrees of consciousness, from the distinct stage of intelligence down to the confused stage of matter. On the contrary mind is mind and Time is Time. Mind exists only on its own level of existence. I mean that in the matrix of all existence, Space-Time, there is an element Time *which performs the same function* in respect of the other element Space as mind performs in respect of its bodily equivalent. *The points of Space have no consciousness in any shape or form, but their instants perform to them the office of consciousness to our brains* .. [Hence,] rather than hold that Time is a form of mind we must say that mind is a form of Time. This second proposition is strictly [as opposed to analogically] true. Out of the time-element .. the quality of mind as well as all lower empirical qualities emerge, and this quality mind belongs to or corresponds to the configuration of time which enters into the space-time configuration which is proper to the level of existence on which mind is found [emphasis added]. (Vol.II, p.44)

Brettschneider (1964) maintains that Alexander's conception of Space-Time is closer to the absolute or Newtonian ("container") view than to the relational view of Space-Time; for Alexander, "things are not only related spatially, but they themselves *occupy* spaces and have spatial forms [emphasis added]." (p.1) However, Alexander (1920) argues for a stronger position, anticipating developments within relativity theory (chapter 4) as to the link between matter, energy and spacetime:

> Another hypothesis as to the connection between things or events and the Space and Time they occupy is that Space and Time are not merely the order of their coexistence or succession, but are, as it were, *the stuff or matrix (or matrices) out of which things or events are made, the medium in which they are precipitated and crystallised*; that the finites are in some sense *complexes* of space and time [emphasis added]. (Vol.I, p.38)

As Brettschneider states, "Alexander would have us abandon the common sense notions of Newtonian physics that construe things and their relations as so much `matter in motion'." (p.63) Alexander maintains that matter is nothing other than "a complex of motion, that is made out of the original stuff which is Space-Time." (Vol. II, p.50) This is of decisive significance since standard cellular automata (CAs), as introduced in chapter 2 and examined further in section 5.5, provide a means by which to implement computational emergentism. CAs map isomorphically onto Newtonian-type dynamical systems and it is at least *conceivable* that a variant of the CA formalism supporting relativistic phenomena can be constructed; consequently, a connection can be established between CA computationalism and Alexanderian metaphysics.

However, there is a distinction between Alexanderian metaphysics and standard CAs with respect to motion: Motion in CAs at the level of `matter' assumes the atomistic or Newtonian form, that is, `particles' changing their locations in space over time. In section 5.5, a variant of the standard CA formalism will be examined which can be used to resolve this difference, thereby establishing a more complete correspondence between Alexanderian metaphysics and CA-computationalism. Another distinction that arises between the two, viz. CAs and Alexanderian Space-Time, is that the former are ontically *discrete* and only *contingently* infinite - as, for example, in Conway's Game of Life (chapter 2) - while the latter is ontically *continuous* and *necessarily* infinite. The assertion of spatio-temporal infinitude is contestable given a `Big Bang' scenario, that is, an *origin* of the physical universe. (One possible solution to this problem involves the adoption of an inflationary cosmology (Linde,94) in which the Big Bang is viewed as merely one `bang' amidst a potential infinity of bangs[14].) With respect to the postulate of a spatio-temporal continuum, Fredkin (1996) states that

> amazingly, there is not, in all of physics, science, or nature, a single case where a basic phenomenon once in doubt as to whether it was continuous or discrete is now known to be continuous. Of course, insofar as the prime quantities of physics, measures of space-time, we can speak of scales that show no deviation from continuous, but no one can claim that we know that space-time physics is continuous, as opposed to discrete, down at more microscopic levels such as at Planck's length. (p.120)

In support of this position, Toffoli (1994b) maintains that

> the problem with differential equations [which is the scientific mode of approach on a continuum view of nature] is that the recipe itself is an infinitesimal one, and has to be executed over a set of points having the infinity of the continuum. It's a task for angels, not for men; we can only carry it out in an approximate way. (p.3)

Consequently, Toffoli, following Fredkin, postulates an atomistic ontology for the natural world, viz. "the ingredients of our physical worlds are discrete particles" (p.3) which is reflected in a computational context by the assertion that "it is differential equations that are the poor man's cellular automata - not the other way around!" (p.4) The continuum postulate is held to emerge as a consequence of *averaging*, viz.

> as soon as the numbers [of objects in a system under consideration] become large enough for averages to be meaningful - say, averages over spacetime volume elements containing thousands of particles and involving thousands of collisions - a definite continuum dynamics emerges. (p.5)

This leads Toffoli to postulate the following:

1. Continuous-looking behaviour is bound to emerge, at a macroscopic scale, from virtually *any* fine-grained mechanism.

2. Virtually all of the differential equations of physics are among those that are known to be limiting

[14] This idea is examined in more detail in chapter 6.

behaviours of simple, discrete fine-grained mechanisms (p.12).

The grounding of an *epistemological* continuum in an ontologically discrete substrate raises an interesting issue introduced in chapter 3: On this view, the physical world is ontologically discrete and finite. The real-number contiuum emerges from averaging over a large but finite number of discrete objects. Since the universe is finite, computation of reals generate numbers with finite expansions. However, if materialism is correct and the physical world is all there is, how can minds *conceive* of (not *actualize* or compute) numbers with infinite expansions ?

Finally, it is worthwhile comparing Alexanderian Space-Time and Einsteinian spacetime as shown in Table 5.3:

Einsteinian spacetime	*Alexanderian Space-Time*
static and finite block universe	dynamic and infinite block universe
event ontology	event (point-instant) ontology
absolute underlying laws	absolute underlying laws
absolute motion undetectable (physically)	absolute motion postulated (metaphysically)
matter and energy distort space; space affects motion of matter and energy	Space-Time is primordial; matter and energy are emergents of Space-Time
forcefields = structural properties of spacetime	forcefields = Space-Time universals
matter = stable regions in forcefield	matter = stable regions in forcefield
spacetime continuous	Space-Time continuous
local field action	local interactions of space-time complexes
no objective, observer-independent reality	absolute Space-Time
arrow of time = result of measurement	arrow of time = ontical; time is ontological and directional
deterministic	deterministic
matter/energy primordial	matter/energy derivative

Table 5.3 Comparison of Einsteinian spacetime and Alexanderian Space-Time.

As stated in chapter 4, Alexander maintains that his conception of Space-Time is consistent with the Einsteinian spacetime block view of the world; hence, Alexander's ontology does not violate relativity theory. However, there is a problem: Einsteinian spacetime is static whereas Alexanderian Space-Time is dynamic. According to the former, it is things, that is, complexes of space-time, which are dynamic; the world considered as a whole is itself static since it is identical to the spacetime block. Furthermore, on the Einsteinian view, matter is primordial and responsible for the curvature of space. According to Alexander, on the other hand, matter and gravitational forces are empirically emergent from Space-Time. (For this reason, Euclidean geometry is an abstraction of Einsteinian spacetime, whereas it is a primordial characteristic of the ontological uniformity of Alexanderian Space-Time; this is important because Euclidean geometry is assumed in the definition of standard CAs[15].)

5.4.3. Categories

Alexander (1920) identifies the following as the ontological categories of Space-Time:

- identity, diversity, existence
- universality, particularity and individuality
- relation
- order
- substance, causality, reciprocity
- quantity and intensity
- whole, part, number
- motion

and maintains that

> [as] to the question whether the *a priori* characters of the world are derived in some manner from experience of things or are primordial and ultimate, the answer is that they are primordial; they do not come into being otherwise than as all things come into being and because things come into being. All things come into being endowed with the categories and with all of them. They are the determinations of all things which arise within Space-Time, which is the matrix of things, `the nurse of becoming'. (Vol.I, pp.330-331)

Thus, Collingwood (1945) is led to maintain that according to Alexander, "space-time is the *source* of the categories, but *they do not apply to space-time*; they belong only to what exists, and what exists is not space-time itself but only the empirical things in it[16]; but these things possess categorial characteristics for one reason and one reason only - namely, that they exist in space-time. Hence Alexander regards them as depending on the nature of space-time: that is, he aims at deducing them from the definition of space-time as its necessary consequences [emphasis added]." (p.162) On Collingwood's reading, Alexander's position is paradoxical: On the one hand, the categories are *a priori*, that is, non-emergent; on the other hand, Space-Time is the source of the categories. Hence, with respect to the primordial stuff, categories are *a posteriori*. The paradox is resolved by contextualizing the senses of emergence: Categories are *a priori* with respect

---

15 Although this is true in principle, this is not the case in fact since the role of geometrical factors is not explicitly addressed in the definition of CAs. An extension to the standard CA formalism, viz. Masked CAs (or MCAs), in which Euclidean geometrical factors are incorporated explicitly is described in (Ali,94a).

16 However, as will be seen in chapter 6, this position is problematic since to the extent that Space-Time is, it must partake of Being and hence, exist (in some sense of the term).

to empirical existents and *a posteriori* with respect to Space-Time; hence, there are directional relations from the ontological substrate (Space-Time) to the categories and from the categories to empirical existents. Relations of this kind will be examined in chapter 6 when the link between Being, beings and categorial `cutting' is investigated.

Brettschneider (1964) maintains that the Alexanderian categories are *a priori* since "they are properties or basic determinations of Space-Time .. which emerge as Space moves along its coefficients of Time." (p.34) (This view is consistent with that of Collingwood (1945) as stated previously.) To reiterate a distinction made in section 5.4.1, categories are ontological, necessary and non-emergent whereas qualities are phenomenal (empirical), contingent and emergent. Only the categories of universality, relation, substantiality, causality and motion will be examined in this section since it is these categories which are of defining significance in connecting Alexanderian metaphysics and CA-computationalism: Universality is important since it relates to the issue of self-organization which in a CA is supported by the necessary existence of attractors (chapter 2); an examination of the Alexanderian category of substance is important because substances are usually held to be ontologically primitive. Alexanderian substances, it will be seen, are equivalent to space-time patterns in CAs; finally, it is necessary to examine the categories of causality and motion since these categories are intuitively primitive and manifested in the context of CA-computationalism as state-transitions. (Only a brief outline of the idea of existence is presented below; the whole-part and existence categories will be examined in greater detail in chapter 6 when the concept of *poiēsis* (or coming-forth) and the link between Being and beings are investigated.)

5.4.3.1. Identity, Diversity, Existence

Brettschneider (1964) maintains that for Alexander, "the *identity* of a thing lies in its particular organization of space-times, and because it occupies a particular set of point-instants, and not others [emphasis added]"; furthermore, "one thing is *different* from another because differences in organization of space-times are occasioned by the occupation of different sets of space-times by different finite entities [emphasis added]." (p.35) According to Alexander, "being is the determinate occupancy of a portion of Space-Time in such a way that the rest of Space-Time is excluded from this portion. Being is .. determinate, spatio-temporal existence. A finite entity approaches the absolute in being as its internal organization increases in all-inclusiveness and harmony." (p.36) Accordingly, existence is closely connected with the category of substance which is discussed below. Importantly, for Alexander, being is not epistemological; empirical existents (or emergents) are real and not merely observational artifacts. This position is consistent with "strong" computationalism in which an ontological version of the coherence theory of truth (chapter 3) is assumed.

5.4.3.2. Universality

According to Brettschneider (1964),

> Plato regarded Space as the matrix of which things are modeled after the image of the Forms. Alexander, however, does not regard form as a property of Space. It is a configuration of Space-Time. Space-Time is not only the stuff of which material things are constituted, but, like the Platonic Forms, it is historically prior to matter. Matter is a finite determination of pure Space-Time. (p.59)

On Alexander's view, the forms are themselves patterns of Space-Time[17]. In place of Platonic forms or ideas *subsisting* in a transcendent realm and only immanent in matter (chapter 2), Alexander advances a strict immanentist position, referring to the pervasive features of space-time complexes as `laws of construction', `plans', `habits' or simply, universals. Furthermore, universality implies both "identity of kind" (Vol.I, p.208) and "a category or determination of Space-Time." (Vol.I, p.214) A superficial reading of *Space, Time and Deity* (1920) presents conflicting positions regarding the category of universals: On the one hand, there are statements in support of Platonic formism (chapter 2) such as "empirical universals are plans of configuration of particulars which are identical in kind. They may be called patterns of configuration or, to use the old Greek word, `forms' of Space-Time. They are essentially in their simplest terms spatio-temporal forms or shapes." (Vol.I, pp.214-215); on the other hand, the *idea* of universals itself is contested. For example, it is maintained that "strictly speaking, there is no such thing as a particular or a universal. All things are individuals." (Vol.I, p.208). The apparent contradiction is resolved once it is appreciated that the categories are "*a priori* plans of configuration" (Vol.I, p.215) and that universality is "begotten like the other categories by Time on Space." (Vol.I, p.217) Alexander clarifies this position as follows:

> the form or configuration of motion belongs not to Space but to Space-Time or motion, and *form does not affect the matter from without*, but belongs intrinsically to any finite piece of Space-Time [emphasis added]. (Vol.II, p.49)

Thus, a commitment to immanentism is implied. Additionally, Alexander is explicit in his rejection of a transcendent Platonic realm of subsisting ideas:

> [On the transcendentalist view,] plans, it may be thought, of space-time are nothing but the universals of different patches of Space-Time, the circular plan, for example, the universal of all circular patches. They are but particular applications of a conceptual universal which is prior to Space and Time and is supplied from understanding or thought, it matters not how. Universality belongs to Space-Time but comes down upon it, either it may be imagined from mind or from some eternal regions as the Forms are supposed to enter Space by Timaeus. Our answer is the old one. *It is not because there are universals that any space-time has a plan, but because Space-Time is uniform .. and admits a plan that existents which are patches of space-time possess universality* [emphasis added]. (Vol.I, p.217)

> [Hence,] the universal subsists in so far as its particulars exist and is spatio-temporal though not particular .. It is not timeless or eternal as being out of time, but as being free from limitation to a particular time. (Vol.I, p.222)

This position is clarified by the statement that "there is no question of any plan [or form]

---

17 However, this position is problematic for a number of reasons which are discussed in section 5.7.

mediating between the particular and the uniformity of Space-Time; the plan is an *embodiment* of that uniformity. The universality of the plan is the capacity of Space-Time to respond on each occasion according to that plan [emphasis added]." (Vol.I, pp.219-220) Furthermore, "particulars are complexes of space-time and belong therefore to the same order or are of the same stuff as the universals which are plans of space-time." (Vol.I, pp.220-221) According to Brettschneider (1964), "the universal is real because it *is* Space-Time and has identifiable patterns or configurations that are discernible because repeated now and again." (p.37) Brettschneider goes on to make the following important point regarding the subsumption relation between universals and particulars, the latter of which are equivalent to organizations of point-instants:

> The relation between particular and universal is not simply a logical subsumption of classes; it is an *historical* relation. A configuration of Space Time, i.e. a determinate finite existent, is a particular. To ascribe universality to a particular is to make an historical judgement, a judgement in fact about the possibility of repetition in Space-Time [emphasis added]. (p.37)

This purely immanentist and historical view of universals is consistent with the interpretation of *emergents*[18] in CAs in which dynamic spatio-temporal organizations are interpreted in contingent, immanentist, non-formalistic and processualist terms (chapter 2).

5.4.3.3. Relation

According to Alexander,

> relations .. are the spatio-temporal connections of things, these things themselves being also in the end spatio-temporal complexes. Since Space-Time is continuous, the connecting situation which constitutes a relation is but spatio-temporal continuity in another form. The relations and the things they relate are equally elements in the one reality and so far are separate realities .. [Things] must at least be connected in Space and Time, and it is plain that they must be connected by all the relations which arise out of the categories, seeing that categories are pervasive features of all things. (Vol.I, p.249)

Brettschneider (1964) offers the following explanation of the distinction between internal and external *relations*[19]:

> Relations are external when outside of and not inherently connected to the terms they come between. On this account, the fact of a relation makes no difference to the terms related. When a term enters into a relation with another term, the connection is accidental to the terms as related. In other words, it matters not to rider, horse, or destination whether I or anybody else ride a cock-horse or anything

---

18 As will be shown in chapters 6 and 7, CA *substrates*, by contrast, are *defined* in formalistic and mechanistic terms. Furthermore, since emergents follow *of necessity* from the definition and intialization of such computational systems, the former can - derivatively - be *interpreted* formalistically.

19 The difference between internal and external relations is examined in more detail in chapter 7.

> else to Banbury Cross or any other place you choose. The realist holds that the range of significance of the variables of a propositional function is limited only by the universe of discourse. *The fact of a relation does not enter into the being of the terms related.* The terms refer to separate, discrete entities. The realist transforms the logical doctrine of external relations into an ontological doctrine. A thing is not necessarily altered when it enters into a relation. When my typewriter rests on top of my desk, its essential being is by no means modified by this relation. In a pluralistic universe [as entailed by certain strains of externalistic realism] each thing is what it is, and cannot both be itself and something else at the same time. In its discreteness it is discontinuous with other existents. Were an entity continuous with other existents not itself, then its intrinsic nature would be altered by its relations [emphasis added]. (pp.44-45)

Brettschneider (1964) offers the following summary of Alexander's metaphysics with respect to the issue of relations:

> Alexander's universe is patterned along lines set forth by the objective idealists. It is a [monistic] block universe of internally connected Space-Time relations. Every element in the universe is connected with every other element. Determinate finite existents are but highly coherent patterns of organization of the primal stuff of which all things are constituted. *All relations are therefore internal.* But some elements are more closely related than others. The distinction between internal and external relations is thus a difference in degree and not in kind. This is the doctrine of degrees of reality [which is fundamental to objective idealism]. Every element of Space-Time is connected with every other element in terms of spatio-temporal perspectives. The Alexanderian universe is thus a unity bound by internal relations. It is not a `seamless' unity, but the universe, taken as a whole in terms of its basic constituents, is one. The basic stuff of pure Space-Time is disparated into different patterns of organization. These are finite events, things, substances, etc. These participate as individuals in either internal or external relations. (pp.49-50)

Thus, relations are both internal and external, depending on whether a holistic or partial perspective is adopted; the internalist perspective is, however, ontological and *a priori* since finites which are externally-related *emerge* from the infinite matrix of Space-Time. It is important to appreciate from the outset that the internal relations in Alexander's Space-Time metaphysics are ontologically different to those appearing in objective idealist schemes such as Whiteheadian panexperientialism (chapter 1). This distinction is linked to the fact that on Alexander's framework, qualities are emergents which appear *ex nihilo* (that is, from nothing[20]) as correlated properties of specific space-time complexes; in panexperientialism, by contrast, experience is ontological, that is, non-emergent and not *merely* empirical (Whitehead,78) (Griffin,98). This point is extremely important since, as has been shown, Alexanderian ontology and CA-computationalism can be closely related. For example, in standard CAs, the state-transition function for each FSM is defined in purely externalist terms (chapters 2, 6 and 7); however, the functional connectivity of the FSMs establishes a functionally- continuous (in the sense of connected) substrate. In what sense, then, can spatio-temporal patterns emerging in a CA be regarded as "separate realities", that is, as discontinuous entities or, in Alexanderian terms, `empirical existents' ? Cariani (1989, 1991) has examined this issue in some detail, maintaining that the delineation of local patterns (finite existents) within

20 The various interpretations of the concept of nothing are examined in chapter 6.

CAs (the Space-Time substrate[21]) involves a shift in observational frames between micro- and macro-dynamic levels of system description: Space-time patterns are identifiable as discrete entities at the *global* (emergent) or macro-level by *observers* situated at that level; these patterns are unobservable at the *local* (substrate) or micro-level. This might appear to contrast with Alexander's position since, as Brettschneider (1964) maintains, "what he calls finite only appears to be so from the standpoint of other finites." (p.54) However, if the *intrinsic* observer perspective (chapter 3) is adopted, then spatio-temporal patterns are discernible only because of the "filtering" capacities of observers who are themselves (on this view) *internal* to the system. Both the intrinsic-emergence or intrinsic-observation perspective within CA-computationalism and Alexander's conception of the link between empirical existents and Total Space-Time are based on an implicit adoption of the ontological interpretation of the coherence theory of truth (or reality), that is, *objective idealism*, and a commitment to a doctrine of reality defined in terms of *internal* relations (chapter 7). As Brettschneider (1964) states, "Alexander's finites are not externally related by any means. They are not the separate, discrete entities that thoroughgoing realists insist upon. They appear finite only when viewed from the standpoint of a limited perspective, not when seen in relation to Total Space-Time, the synthesis of all perspectives." (p.55) However, as shown in section 5.7, CAs at the ontological or component (FSM) level *are* externally-related systems[22].

5.4.3.4. Substance

According to Alexander,

> all existents, being complexes of space-time, are substances, because any portion of Space is temporal or is the theatre of succession [and] all succession is spread out in space. (Vol.I, p.269)

Furthermore,

> qualities .. are correlated with certain motions; and it is indifferent for our purpose whether the quality belongs, as will be here maintained, to the motion itself; or belongs to mind and is the mental correlate of the motion, as is the belief of those who distinguish primary and secondary qualities, but recognize a primary correlate of the secondary quality. A thing or complex substance is then a contour of space (i.e. a volume with a contour) within which take place the motions correlated to the qualities of the thing; and the complex substance or thing is the persistence in time of this spatial contour with its defining motions. (Vol.I, p.270)

> The persistence of a piece of Space in Time which results from the retention of the configuration of its movements according to its law of construction does not of course imply that the piece of Space

---

21 It must be appreciated from the outset that the spatio-temporal substrate in a CA is not necessarily infinite; hence, CA Space-Time is not necessarily fully isomorphic with Alexanderian Space-Time (which is infinite *by definition*).

22 As stated previously, this position is explained further in chapters 6 and 7.

> is stationary as a whole. On the contrary, no substance occupies the same place continuously, if only because of the movement of the earth or other heavenly body, and it may change its place also by locomotion or transference. But the contour and internal configuration remain within limits the same, though not the position of the whole thing. (Vol.I, p.271)

Brettschneider (1964) maintains that "in Alexander's architectonic design, substance evolves out of an infinite with which the finite is continuous, not discontinuous[23] .. Total Space-Time is a continuous whole. It breaks up into finite substances through the agency of the coherence of spatial elements changing temporal coefficients." (p.53) Since Alexander does not allow for a Platonic realm of forms, it is necessary for him to postulate Time as the creative principle which gives rise to substances (empirical existents) through its own irreducible action. In certain CAs, by contrast, creative self-organization is a consequence of the existence of attractors which are themselves necessary consequences of the *irreversibility* (section 5.5.5) of the local FSM state-transition function in the CA. As stated previously, and consistent with Alexander's scheme, CA-substances or empirical existents are identifiable with the dyamic spatio-temporal patterns formed by groups of cells (FSMs) in specific states.

5.4.3.5. Causality

According to Alexander, "there is no causality in the continuance without change of the same motion .. a motion does not cause its own continuance, is not as it were the *cause* of itself, but *is* itself [emphasis added]." (Vol.I, pp.281-282) For Alexander, cause and effect must, therefore, be *different*, that is, result in different motions. In addition, causes precede effects; on his scheme, temporally-retroactive causation (chapter 3) is disallowed due to the asymmetrical `forward-directionality' of Time. Furthermore, Alexander maintains that

> the only self-contained reality in which all causality is immanent [or *internally*-affective; that is, causality is between the substances *within* a complex substance] is the universe itself, and its immanent causality is but the transeunt [that is, *externally*-affective] causality of the existents it contains. (Vol.I, p.284)

Since Alexander adopts an ontological interpretation of the coherence theory of truth, he is led to assert that "there is no causal relation between the infinite whole and any one of its parts. There is only such relation between one part and another." (Vol.I, p.288) The interpretation of causality that this view necessitates will be examined further in chapter 6. However, it should be sufficient to note here that this view is consistent with CA-computationalism; causation occurs "bottom up" in CAs, viz. at the ontological or substrate level of functionally-connected FSMs.

---

23 In this connection, Alexanderian Space-Time is similar to Heideggerian Being (chapters 1 and 6).

5.4.3.6. Motion

On Alexander's view,

> it might be objected that a motion or a bit of Space-Time is a really existent concrete thing and therefore cannot be a category. Such an objection would imply a complete misunderstanding of the nature of categories. They are not expressing mere adjectives of things, but concrete determinations of every space-time. Existence is the occupation of any space-time. (Vol.I, pp.320-321)

Thus, for Alexander, motion is both a category and an empirical existent, that is, a space-time complex. This position is explicitly asserted in the follows statement, viz.

> as to motion it is to be described indifferently as empirical or categorial, for it is the meeting-point of the two. (Vol.II, p.67)

This is, in fact, the case for all the categories as stated previously in section 5.4.3. Furthermore, Alexander maintains that

> point-instants are real but their separateness from one another is conceptual. They are in fact the elements of motion and in their reality are inseparable from the universe of motion; they are elements in a continuum. So far from being finites, they are the constituents which are arrived at as the result of infinite division and belong to the same order as the infinites. Consequently they must be regarded not as physical elements like the electrons, but as metaphysical elements, as being the elementary constituents of Space-Time or Motion. (Vol.I, p.325)

Hence, "movement is anterior to things which are complexes of movements, and it is quite true that that movement is a stuff of which things are made and this is not a mere relation between things which already exist and are said to move." (Vol.I, p.329) In the context of CA-computationalism, the primordiality of motion is substituted by the primordiality of computation which is manifested at the ontological (or substrate) level in the state-transitions of the functionally-connected FSMs which realize the CA.

5.4.4. Emergence

It was stated in section 5.3.4 and has been restated throughout section 5.4 that emergence is a foundational concept in Alexander's metaphysics. Although the categories associated with Space-Time are non-emergent as appearing in finite empirical existents (phenomena), according to Alexander, the (secondary) *qualities* associated with specific space-time complexes are emergent, viz.

> the emergence of a new quality from any level of existence means that at that level there comes into being a certain constellation or collocation of the motions belonging to that level, and possessing the quality appropriate to it, and this collocation possesses a new quality distinctive of the higher complex. (Vol.II, p.45)

Although "ascent takes place, it would seem through complexity .. at each change of quality the complexity as it were gathers itself together and is expressed in a new simplicity." (Vol.II, p.70); consequently, as Brettscheider (1964) states, "Alexander thinks of the different qualities of existence as forming a discontinuous series." (p.29) On Alexander's view, "life is not a consciousness with something of its powers left out, nor materiality consciousness with still larger omissions and imperfections. The difference is one of kind or quality and not of degree." (Vol.II, p.69)

Alexander's concept of emergence can be contrasted with that due to Bunge (1979a). It can be argued that the former adopts what may be described as a *cumulative* or `conservative' approach to emergence: As Brettschneider (1964) states, "the emergence of a new quality refers to the rearrangement of a particular complex of Space-Time at a given level of existence in such a way that a new pattern develops *in addition to* the older finite features [emphasis added]." (p.57) Consequently, it could be asserted that on Alexander's scheme, the mental is *also* neural, biological, physico-chemical and spatio-temporal and such that none of the properties associated with lower order phenomena are displaced following the emergence of the new phenomenon with its associated properties. Bunge (1979) contests this position maintaining that preservation of properties does not constitute a *necessary* condition for emergence, viz.

> the breakdown (dismantling) of a system, and the substitution of some of its components, are emergence processes .. Every assembly process is accompanied by the emergence of some properties and the loss of others. I.e. let the parts of a thing *x* self-assemble into a system during the interval [*t*,*t*']. Then the system lacks some of the properties of its precursors - i.e. $p_x(t)-p_x(t')\neq\emptyset$ - but on the other hand it possesses some new properties - i.e. $p_x(t')-p_x(t)\neq\emptyset$. (p.30)

However, Alexander's position is, in fact, consistent with both Brettschneider's interpretation of it and Bunge's alternative since as Alexander states, "the empirical qualities of the `material' are carried up into the *body* of the higher level but *not* into its new quality [emphasis added]." (Vol.II, p.70) (On Alexander's scheme, the dualistic abstractions of Space and Time are analogous to body and mind; thus, property-*preservation* occurs at the empirical substrate level while simultaneously property-*loss* occurs at the qualitative level of the emergent.) This interpretation of his position is supported by the following statements:

> The body or stuff of each new quality or type of soul has itself already its own type of soul, and ultimately the body of everything is a piece of Space-Time, the time of which is the soul-constituent which is identical with the body-constituent. (Vol.II, p.69)

> A complex of processes on a level *L* with the distinctive quality *l* becomes endowed, within the whole *L*-thing or body, with a quality *l'* and the whole thing characterised by this quality rises to the level *L'*. The processes with the emergent quality *l'* constitute the soul or mind of a thing or body which is on the level *L*. The mind of a thing is thus equivalent only to a portion of that thing .. Thus the soul of each level is the soul of a body which is the stuff of which it may be called the form. (Vol.II, p.68)

Thus far in the discussion it has been stated that Alexander's scheme is emergentist: For example, and according to Collingwood (1945), on Alexander's ontology, the world of nature appears "as a single cosmic process in which there emerge, as it goes on, higher orders of being." (p.158) What has not yet been addressed is *why* emergence should occur at all. In short, what causes the emergence of finite quality-bearing empirical existents ? Alexander maintains that

> empirical things or existents are .. groupings within Space-Time, that is, they are complexes of pure events or motions in various degrees of complexity. Such finites have all the categorial characters, that is, all the fundamental features which flow from the nature of any space-time, in an empirical form - each finite has its proper extension and duration, is built on the pattern of its specific universal, in a substance of a certain sort and the like. (Vol.II, p.45)

> New orders of finites come into existence in Time; the world actually or historically develops from its first or elementary condition of Space-Time, which possesses no quality except .. the spatio-temporal quality of motion. But as in the course of Time new complexity of motions comes into existence, a new quality emerges, that is, a new complex possesses *as a matter of observed empirical fact* a new or emergent quality [emphasis added]. (Vol.II, p.45)

It would appear from this last statement that there is no reason (that is, *necessity*) for emergence, it simply (and *contingently*) happens; that is, emergence is merely an "empirical fact". This view is supported by the following statement, viz.

> the higher quality emerges from the lower level of existence and has its roots therein, but it emerges therefrom, and it does not belong to that lower level, but constitutes its possessor a new order of existent with its special laws of behaviour. The existence of emergent qualities thus described is something to be noted, as some would say, under the compulsion of brute empirical fact, or, as I should prefer to say in less harsh terms, to be accepted with the `natural piety' of the investigator. *It admits no explanation* [emphasis added]. (Vol.II, pp.46-47)

However, this interpretation of Alexander's position is incomplete as the following statement establishes:

> Empirical things come into existence, because Space-Time of its own nature breaks up into finites, the lowest such finites being simple motions of different velocities or intensities of motion and different extents of it. *Time and Space, either of them, creates differences in the other or breaks it up*. But in a special sense Time is the author of finitude, for it is the transition intrinsic to Time which in the first place makes motion possible, and secondly provides for the ceaseless rearrangements in Space through which groupings of motions are possible. Time could not do its work without Space; but, this being presumed, *Time is the principle of motion and change*. (Vol.II, pp.47-48)

Thus, according to Alexander, it is the restlessness of Time, which is the source of all movement and motion in the world[24], that is responsible for the creativity manifested in emergent evolution, in short, for *cosmogenesis* (Brettschneider,64): "Time is the ordering [or coherence] quality of the extended structure that is Space." (p.73) Importantly,

---

[24] In this connection, it is interesting to note that on Heideggerian ontology, Temporality (or time) is the `horizon' for the interpretation of Being (or existence). However, as will be seen in chapter 6, Heideggerian Being, in contrast to Alexanderian existence with its grounding in a Space-Time event monism, is *radically* pluralistic.

change is not regarded as categorial (that is, ontological[25]) since "change always implies movement and is movement from one movement to another. Change is an alteration in something else, viz. in movement." (Vol.I, p.330) However, detailed examination of this position by Brettschneider (1964) reveals the following problems:

> the functions of ordering and creating [of finite empirical existences from infinite Space-Time] require an agent of efficient causation. Alexander's Spinozistic tendencies require that he look within the system for this efficient agent. He is confronted by three possibilities: (1) Space-Time may act upon itself; in efficient causation, however, there must be some differentiation between that which effects a change and that upon which a change is effected[26]; (2) Space may act upon Time; or (3) Time may act upon Space. (p.21)

Alexander adopts the third alternative. However, as Brettschneider points out, this contradicts his assertion that Space and Time are co-equal `abstractions' of the one Space-Time continuum (section 5.4.2). Furthermore, "Alexander's organizing relations must somehow be imported into Space-Time's manner of functioning, since Space-Time is an irreducible simple. Alexander accomplishes this by making motion a function of time. This is what he means by the inherent restlessness of Space-Time." (p.21) This is analogous to the temporal actualization of state-transitions in the basin of attraction field characterizing the behaviour of a (non-reversible) CA. Hence, in both Alexanderian metaphysics and CA-computationalism, the ordering or coherence principle is ontological, that is, intrinsic to the system itself; consequently, both schemes are necessarily committed *a priori* to the idea of emergence as self-organization (chapter 3). The validity of the application of this concept in natural and artificial contexts will be critically examined in chapter 6 when the interpretation of emergence as *poiēsis* (coming-forth, bringing-forth, becoming) is examined.

In this study, three phenomenal levels (matter, life and mind) have been identified and briefly described (chapter 4), phenomena which provide the means by which to examine the possibility of a unification of computationally emergent artificiality. This selection is very close to that made by Alexander (1920) himself, viz.

> roughly speaking, the different levels of existence which are more obviously distinguishable are motions, matter as physical (or mechanical), matter with secondary qualities, life, mind. (Vol.II, p.52)

The scheme outlined in section 5.5 identifies motions, matter (bearing both `primary' or mechanical and `secondary' or experiential qualities), life and mind. In both Alexander's

---

25 This marks yet another point of distinction between Alexanderian and Heideggerian ontology: As will be seen in chapter 6, on the latter, categories (*Kategoria*) are *hypostatizations* (static closures) of the *logos* which reflects an essentially *dynamic* openness to categorial (more precisely, modally existential) emergence in Being.

26 Hence, the problem of *self*-organization discussed in chapter 3.

ontology and CA-computationalism, motions are viewed as categorial features of and empirical emergents from Space-Time; on CA-computationalism, motions are identified with computations (that is, programs in execution or processes). Before this computationalist scheme is outlined, Alexander's conception of matter, life and mind will be briefly examined.

5.4.5. Matter

The phenomenon of matter was discussed implicitly during the examination of ontological Space-Time and categoral motion in sections 5.4.2 and 5.4.3 respectively. However, Collingwood (1945) provides the following summary of the Alexanderian conception which serves to further explicate many of the underlying issues involved:

> everything that exists has a place-aspect and also a time-aspect. In its place-aspect it has a determinate situation; in its time-aspect it is always moving to a new situation; and thus Alexander arrives metaphysically at the modern conception of matter as inherently possessed of motion, and of all movements as relative to each other within space-time as a whole. The first emergence is the emergence of matter itself from point-instants: a particle of matter is a moving pattern of point-instants, and because this is always a determinate pattern it will have a determinate quality. (p.160)

As Brettschneider (1964) states,

> matter is the first level of emergence. Specific organization patterns of Space-Time are denoted by matter; that is, matter is finite, the simplest type of construct to emerge out of infinite Space-Time's primordial nature. (p.61)

Alexander identifies inertia or mass and energy as characteristic empirical existents at the level of matter. In contrast to proponents of mechanism who, taking their lead from Hume, distinguish between primary and secondary *qualities* (chapter 2), he maintains that "shape, size and motion and number (the traditional primary qualities) are not qualities at all. They are [*categorial*] determinations of the thing, but are misnamed qualities because the secondary characters, colour, temperature, taste, and the like, are qualities, and the primary features are ranged into one class with them as a contrasting group within the class." (Vol.II, p.56) For Alexander, "the secondary quality is the mind or soul of its corresponding vibration or whatever the primary movement may be." (Vol.II, p.59) This leads him to endorse a generic version of the identity theory (chapter 4) in the context of the emergence of empirical phenomena:

> [the] secondary qualities are thus a set of new qualities which movements of a certain order of complexity have taken on, or which emerge with them; and the material movements so complicated can no more be separated from the secondary quality (which is not merely correlated with them but identical with them) than the physiological processes which are also psychical can be what they are in the absence of the conscious quality. (Vol.II, p.59)

Anticipating ideas already discussed in connection with an examination of a unifying teleological framework presented by Campbell (section 5.2.2), Alexander speculates as to whether "in these ages of simpler [material] existence something corresponding to the method pursued by nature in its higher stages, of natural selection [was in operation] .. whether that is to say, nature or Space-Time did not try various complexes of simple motions and out of the chaos of motion preserve certain types." (Vol.II, p.55) As stated previously (section 5.3.2), it may be that the notion of self-organization, which is, in fact, implicit in Alexander's commitment to a coherence-theoretical position (section 5.4.4), is a more appropriate teleological concept for matter.

5.4.6. Life

Alexander is emphatic in stating that "life is not an epiphenomenon of matter but an emergent from it." (Vol.II, p.64) Furthermore, and importantly, he maintains, against the vitalists, that "the directing [or teleological] agency is not a separate existence but is found in the principle or plan of the constellation." (Vol.II, p.64) (This supports the link between a coherence-theoretical position and self-organization stated in section 5.4.5.) As Collingwood (1945) states, for Alexander

> living organisms .. are patterns whose elements are bits of matter. In themselves these bits of matter are inorganic; it is only the whole pattern which they compose which is alive, and its life is the time-aspect or rhythmic process of its material parts. (p.160)

This position is clarified by Alexander himself as follows:

> Life [is] an emergent quality taken on by a complex of physico-chemical processes belonging to the material level, these processes taking place in a structure of a certain order of complexity, of which the processes are the functions. A living process is therefore also a physico-chemical one; but not all physico-chemical processes are vital, just as every mental process is also physiological but not all physiological ones are mental. (Vol.II, pp.61-62)

Alexander does not examine the conditions necessary for something to be classified as living in any detail; on his view, organization (or complexity) is necessary but not sufficient for life (additional necessary characteristics including self-regulation and self-reproduction). However, from the above statements, it should be clear that his formulation of the nature of the living is consistent with the computational theory of life (CTL) and the possibility of artificial life or A-Life (chapter 4). Statements such as the following indicate implicit antecedent support for the computationalist thesis:

> If the study of life is not one with a peculiar subject-matter, though that subject-matter is *resoluble without residue* into physico-chemical processes, then we should be compelled ultimately to declare not only psychology to be a department of physiology, and physiology of physics and chemistry, but, if we are consistent, to be a chapter, like all the other sciences, of *mathematics*, which deals with motion and Space and Time [emphasis added]. (Vol.II, p.63)

Assuming the computational theory of matter (CTMa) described in chapter 4, the following statements provide upport for "strong" A-Life (that is, realization as opposed to mere simulation), viz.

> if we regard the organism as behaving according to the laws determined by its own peculiar structure, a material machine may, since it also obeys the laws of its structure, be said to be alive .. The difference of the material and the organic `machine' lies in the comparative rigidity of the one and the plasticity of the other. Plasticity is not realised by matter but waits for life. But *if we could secure the right sort of machine it would be an organism and cease to be a material machine*. We have no right therefore to confuse the definiteness of mechanism with its materiality, and on this ground cut off the continuity between the material structure and the emergent order of vital structure. The true antithesis is that of the vital and the material and not of the vital and mechanical [emphasis added]. (Vol.II, p.66)

5.4.7. Mind

According to Collingwood (1945), on Alexander's scheme "mind is a pattern of vital activities." (p.161) Consistent with his commitment to an emergentist position, Alexander maintains that "every object we know is a fragment from an infinite whole [Space-Time], and every act of mind is correspondingly a fragment out of a larger though finite mass." (Vol.I, p.23) However, while it is true that "empirical things are complexes of space-time with their qualities" (Vol.II, p.3), Alexander maintains that

> the nature of mind and its relation to body is a simpler problem in itself than the relation of lower qualities of existence to their inferior basis. (Vol.II, p.3)

Consistent with his view of matter (section 5.4.5), he offers an identity-theoretical position with respect to mind:

> We are forced .. to go beyond the mere correlation of the mental with [associated] neural processes and to identify them. There is but one process which being of a specific complexity, has the quality of consciousness; the term complexity being used to include not merely complexity in structure or constitution of the various motions engaged, but also intensity, and above all unimpeded outlet, that is, connection with the other processes or structures with which the process in question is organized. (Vol.II, p.5)

However, not all neural processes are held to be mental on his view, viz.

> while every mental process is *also* neural, it is not *merely* neural, and therefore also not merely vital .. assuming that the conception of localization of mental functions in specific regions of the brain is physiologically correct, we may safely regard locality of the mental process as what chiefly makes it mental as distinct from merely neural, or what distinguishes the different sorts of mental processes from one another. (Vol.II, p.6)

> .. without the specific physiological or vital constellation there is no mind. All less complex vital constellations remain purely vital. Thus not all vital processes are mental. (Vol.II, p.7)

> Mind is identical with some physical counterpart and is connected by some physical connections which need not necessarily be themselves mental ones, carrying the mental quality. (Vol.II, p.25)

Two elements of Alexander's conception of mind are of particular interest in the context of this study, viz. consciousness and determinism.

5.4.7.1. Consciousness

Alexander maintains that what Chalmers (1995, 1996) has referred to as the "hard problem" of consciousness, viz. how subjectivity (the mind) can emerge from an objective substrate (the brain), does not constitute a problem since mind is merely an emergent empirical fact which must be accepted in a spirit of `natural piety':

> No physiological constellation explains for us why it should be mind. But at the same time, being thus new, mind is through its physiological character continuous with the neural processes which are not mental. It is not something distinct and broken off from them, but it has its roots or foundations in all the rest of the nervous system. It is in this sense that mind and mental processes are vital but not merely vital. (Vol.II, p.8)

Alexander rejects physicalist versions of *conscious inessentialism* (Flanagan,95), a thesis which asserts the logical possibility of non-conscious `zombies' that are behaviourally and physically indistinguishable from their conscious counterparts:

> The mental state is the epiphenomenon of the neural process. But of what process ? Of its own neural process. But that process possesses the mental character, and there is no evidence to show that it would possess its specific neural character if it were not also mental. On the contrary, we find that neural processes which are not mental are not of the same neural order as those which are. A neural process does not cease to be mental and remain in all respects the same neural process as before. (Vol.II, pp.8-9)

Searle (1992) maintains that "consciousness .. is a biological feature of human and certain animal brains. It is *caused* by neurobiological processes and is as much a part of the natural biological order as any other biological features such as photosynthesis, digestion, or mitosis [emphasis added]." (p.90) Alexander appears to hold a slightly different view: Brains do not *cause* minds; rather, minds are the empirical qualities *associated* with certain neurophysiological states[27]. However, Searle's position is indeed identical with this position although it serves to clarify that due to Alexander since it is explicit about the conditions under which mind emerges, viz. specific neurophysiological organization. Evidence supporting an identification of Searle's and Alexander's positions can be found in (Searle,92):

> Consciousness is not a `stuff', it is a *feature* or *property* [cf. Alexander's *empirical quality*] of the brain in the sense, for example, that liquidity is a feature of water.
>
> There is no `link' between consciousness and the brain, any more than there is a link between the liquidity of water and H_2O molecules. If consciousness is a higher-level feature of the brain, then there cannot be a link between the feature and the system of which it is a feature. (p.105)
>
> [However,] consciousness *is* a causally emergent property of systems [emphasis added]. (p.112)

27 Hence, Alexander appears to be committed to some version of supervenience (chapter 3).

Crucially, both views assume *causal reductionism* which Searle defines as follows:

> This is a relation between any two types of things that can have causal powers, where the existence and a fortiori the causal powers of the reduced entity are shown to be entirely explainable in terms of the causal powers of the reducing phenomena. (p.114)

However, Searle maintains that causal reduction does not necessarily entail ontological reduction (chapter 3); specifically, a causal reduction of mental processes to neurophysiological processes does not entail an ontological reduction of subjectivity to an objective substrate. Thus, Searle is forced to adopt with Bunge (1977b, 1979a) a form of *ontological pluralism*. However, for Alexander, who is committed to a Space-Time monism, it is not ontological pluralism which is necessitated but *qualitative or phenomenal pluralism*. According to Alexander, the ontological categories constitute a finite set and quality is not included in this set; qualities are merely the emergent `mental' aspects associated with specific empirical complexes of space-time. Tallis (1994) has criticized Searle's position maintaining that it is nonsensical to hold that mental states are both realized in and caused by brain states. Searle's position is based on the view that causality can legitimately occur between different levels in a phenomenal system. However, if levels are merely descriptive (or epistemological) devices, then Searle's position reduces to the notion that causality occurs between different levels of *description* (or syntax). But this would be problematic since Searle contests (1) the validity of the *representationalist* position within cognitive science and the philosophy of mind (Searle,80) (Searle,84) (Searle,92) and (2) the view that syntax (that is, descriptive or representational structure) is *causal*[28]. Thus, it would appear that - suitably interpreted - Alexander's formulation is, after all, more accurate; mental states are not caused by certain brain states, but are *identical* with brain states although not *merely* identical with such states since brain states which are also mental possess a quality not possessed by brain states which are not mental, viz. *experientiality* (ontological-subjectivity, first-personhood or *what-it-is-like-ness*). Mental states are not causal but emergent; causality occurs only at the ontological (or substrate) level of Space-Time (despite Alexander's arguments to the contrary). The quality of mentality is not caused, but emerges *ex nihilo* (chapter 6). As stated previously, according to Alexander, consciousness must be accepted as an empirical (qualitative) fact in a spirit of `natural piety'. Finally, Searle (1992) maintains that "the ontology of the unconscious is strictly

---

[28] Searle (1992) describes the computational theory of mind (CTMi) - or "Strong AI" - as follows: "The thesis is that there are a whole lot of symbols being manipulated in the brain, 0's and 1's flashing through the brain at lightning speed and invisible not only to the naked eye but even to the most powerful electron microscope, and it is these that cause cognition. But the difficulty is that the 0's and 1's as such have no causal powers other than those of the implementing medium because the program has no real existence, no ontology, beyond that of the implementing medium. Physically speaking, there is no such thing as a separate `program level.'" (p.215) This latter position is consistent with the argument presented in chapters 6 and 7 to the effect that relative to *becoming*, computation is ontologically-ontical (that is, defined in terms of the causality of the implmenting substrate), and relative to *Being*, computation is ontically-ontological in that it *prosthetically-extends* the intentional capacities of the human *user*, viz. IA or intelligence amplification (Brooks,94a).

the ontology of a neurophysiology capable of generating the conscious." (p.172) This position is anticipated by Alexander, who identifies the unconscious with neurophysiological complexes `awaiting completion' so as to enter into consciousness. On his view, "[mental processes, memories, dispositions etc] would thus form a permanent undercurrent of the mental life, but would remain purely physiological till called upon to enter into the psychical neural constellation." (Vol.II, p.28) Hence,

> in the absence of the completing conditions which evoke consciousness, the mind slips into a physiological or psycho-physical disposition, which is only potentially conscious, but is actually unconscious. (Vol.II, pp.60-61)

5.4.7.2. Determinism

Alexander advances a form of compatibilism[29] with respect to freedom and volition. He maintains that the conflict between qualitative mental freedom and bodily determinism is merely *apparent*. As he states,

> choice between two alternatives *seems* at first sight to distinguish completely between voluntary choice and ordinary physical causality. For when two forces are operative upon a physical body the effect is the resultant of the two effects of the separate causes; whereas in choosing, one or the other motive is adopted and the other disregarded [emphasis added]. (Vol.II,p.321)

However, according to Alexander

> freedom is nothing but the form which [deterministic] causal action assumes when both cause and effect are enjoyed [that is, subjectively experienced] (Vol.II,p.315)

Hence, "freedom does not mean indetermination." (Vol.II,p.330) Alexander goes on to state that "there is nothing in free mental action which is incompatible with thorough determinism. Neither is such determinism incompatible with novelty." (Vol.II,p.323). According to Alexander, "determinism and *prediction* are .. distinct ideas, and determinism is compatible with unpredictability and freedom with predictability [emphasis added]." (Vol.II,p.329) Hence,

> not only may mental action be determined and yet unpredictable, it may be free and yet necessary [since] the necessity that the will obeys is the `necessity' of causation, the determinate sequence of event upon its conditions. (Vol.II,p.329)

Alexander is committed to a form of epistemological indeterminism which allows for what Davies (1992) has referred to as `deterministic randomness'. As the former states,

---

29 The Oxford Companion to Philosophy (1995) defines *compatibilism* as " a view about determinism and freedom that claims we are sometimes morally dree and responsible even though all events are causally determined."

> the determinism of the free act means no more than this, that is has followed in fact from its antecedents, as they exist in the character of the agent and the circumstances which appeal to him for action. The freedom consists in the act of choice; there is no power of choosing behind the choice itself, no freedom *of* choice but only freedom experienced *in* choice. (Vol.II,p.330)

Hence, for Alexander, freedom is epiphenomenonal, the qualitative correlate of a specific empirical existent; it is causal only to the extent that the spatio-temporal complex with which it is identical is causal and all causation is ontologically bottom-up *from* the Space-Time substrate *to* the emergent phenomenon.

5.4.8. From Space-Time to CA-Computationalism

As shown above, Alexander's ontology provides a suitable foundation upon which to base an examination of the links between computationalism, emergence and artificiality: Space-time *events* are prior to matter, life, and mind. If such phenomena are regarded as concrete then Space-Time must be viewed as abstract (Brettschneider,64). However, computation is also abstract. Hence, the possibility of reinterpreting Alexander's ontology in computational terms by establishing an isomorphism between computations and Space-Time events. However, a computationalism which maps onto Alexander's Space-Time monism is not readily realized (or visualized) in terms of Turing Machines (chapter 2). An alternative realization of computationalism supporting structural as well as behavioural (and functional) isomorphisms with a Space-Time event ontology is to be found in cellular automata or CAs (chapter 2). As Toffoli and Margolus (1990) state, "cellular automata are more expressive than Turing machines, insofar as they provide explicit means for modeling *parallel* computation on a spacetime background." (p.229) Consequently, cellular automata have been adopted as the means by which to realize computationalism in this study. In the following section, a unified framework of emergent artificiality grounded in CA-computationalism is outlined.

5.5. Computationally Emergent Artificiality (CEA)

In this section, a unified framework of emergent artificiality (A-Physics, A-Life, AI) grounded in a CA substrate (CEA) and isomorphic with Alexander's emergent phenomenal hierarchy of naturality (section 5.4) is described. In the discussion that follows, the CA analogue of physics and matter, viz. *CA-matter*, a CA-based realization of the computational theory of matter or CTMa (chapter 4), is examined in detail. There are (at least) three reasons why this is necessary: (1) CA-matter denotes the first emergent phenomenal level in the artificiality hierarchy; (2) computational theories of life (CTL) and mind (CTMi) were examined in chapter 4 whereas discussion of the CTMa was postponed; (3) CA-life and CA-mind - that is, CA-based realizations of the computational theories of life and mind respectively - can be linked directly to notions such as computation- and construction-universality (chapter 2), concepts which have been extensively studied and widely reported in the literature. However, perhaps the most important reason for examining the link between CA-matter and CA-computationalism is implicit in the following assertion of Davies (1992), viz.

> the laws of physics act as the 'ground of being' of the universe. (p.73)

In this study, both ontological and epistemological issues associated with computationalism are being examined. If physics is the ground of being, that is, if physics is ontological then it becomes necessary to examine the link between CA-computationalism and physics so as to determine whether the former can in fact subsume the latter, thereby rendering it derivative and phenomenal as opposed to ontological. If this is the case, then CA-computationalism can replace physics as the ontological ground of the phenomenal universe (Steinhart,98).

5.5.1. Cellular Automata (CAs) Reviewed

The basic characteristics of cellular automata (CAs) were presented in chapter 2; hence, in the following sections only extensions to the standard model are described. However, before these extensions are discussed it is worthwhile reviewing a few of the properties of standard CAs relevant to the issues under consideration.

5.5.1.1. Finitude

Standard CAs are finite in at least two senses: First, as Toffoli and Margolus (1990) state, CA `laws' are *finitary*, that is, "by means of the local [FSM state-transition] map one can explicitly construct *in an exact way* the forward evolution of an arbitrarily large portion of a cellular automaton through an arbitrary length of time, all by finite means." (p.230); second, spacetime is assumed to be discrete with a finite amount of information (state) contained in a finite volume of spacetime (cell). Although this postulate of the finite nature thesis (chapter 2) is adopted in this study, and hence, only digital or *discrete* CA are considered, non-discrete CA have been reported in the literature; for example, MacLennan (1990) presents a variant of the Game of Life (chapter 2 and section 5.5.2) based on a continuous spatial automaton, that is, a CA in which both cells and possible cell states are continuous entities. The other postulate of the finite nature thesis, viz. that the universe is itself a large but finite automaton is not a necessary assumption within standard CAs; for example, the Game of Life was originally conceived in the context of an infinite lattice.

5.5.1.2. Universality

Fredkin (1990) defines a universal cellular automaton (UCA) as a machine that can exhibit any (and every) kind of mechanistic, locally finite behaviour. As he states,

> if you can imagine a process that could take place in a particular CA of any degree of complexity, then the same process can also be done by any CA that happens to be *universal* even though the universal CA is governed by a drastically simpler rule than the complex CA. (p.255)

This is important since according to Fredkin,

> if two processes are identical except for a coordinate transformation, they must be the same process. (p.264)

Thus, "if any particular CA is a good model of microscopic processes in physics, then every UCA could be programmed to exhibit the isomorphic *behaviour* (after a space-time mapping) [emphasis added]." (p.255) This makes possible the *embedding* of machines (CAs) within machines leading to the production - via design or emergence - of a virtual machine hierarchy (chapters 2 and 4). As will be shown in sections 5.5.5 and 5.5.6 when techniques for converting irreversible or *non-reversible* CAs (or NRCAs) into reversible CAs (or RCAs) are presented, it is logically possible to embed non-reversible UCAs (or NRUCAs) in reversible UCAs (or RUCAs) and vice-versa, a consequence of them both being members of the class of UCAs. Consequently, Fredkin maintains that "it is even feasible to use a non-reversible model to model a reversible process; it is just aesthetically obnoxious." (p.257) Additionally, Fredkin argues for a connection between physics and universality, viz.

> if microscopic physics (assuming finite nature) was not universal, then it would be tautologically true that the construction of an ordinary computer would not be possible; but nature allows us to construct computers! (p.257)

Taken generally, the above statement implies that a phenomenon must already exist as a *potentiality* in a supporting substrate if the emergence of that phenomenon from within that substrate is to be possible. This point is extremely important since it entails a reinterpretation of the standard view of emergence (chapter 3) as will be shown in chapters 6 and 7. However, Toffoli and Margolus (1987) implicitly subscribe to the conventional interpretation of emergence in terms of *creatio ex nihilo* (chapter 3). As they state,

> it is often too easy to arrive at models that display the expected phenomenology just because the outward symptoms themselves, rather than some deeper internal reasons, have been directly programmed in .. We want models that talk back to us, models that have a mind of their own. *We want to get out of our models more than we have put in* [emphasis added]. (p.142)

5.5.1.3. Emergence

Computational emergence, emergent computation and the phenomenon of emergence in CAs were discussed in chapter 3. As stated earlier in that chapter, a possible consequence of emergent evolution is that the parts in the new whole may be modified. To recap, a set of elements $\{a,b,c,d\}$ might interact to form an emergent whole X; however, when X is analyzed into its components, these may be $\{e,f,g,h\}$ where $\{a,b,c,d\}$ is not isomorphic with $\{e,f,g,h\}$. The original components have been modified during construction of the whole, viz. the emergent or systemic phenomenon, implying that the structural components used to construct an emergent whole can act as `scaffolding' which is not discarded once the whole has been constructed, but is reintegrated into the whole in modified form; alternatively, the component structures may be retained in the new whole without modification or discarded completely[30]. The three possibilities, viz. (1) component redefinition, (2) component retention, and (3) component disintegration can be demonstrated in the context of CAs if the components (elements) are identified with CA spacetime *complexes* (patterns). In contrast, primitive CA events (state-transitions) are atomic and ontologically static; at this (substrate) level, $\{a,b,c,d\} \equiv \{e,f,g,h\}$.

5.5.1.4. Hierarchies

Adamides et al. (1992) describe a variant of standard CAs, viz. hierarchical cellular automata (HCAs), which they define as "structures of simple CA that transfer state information in one direction at a time, from a specific level of hierarchy to a higher or lower one, depending on the definition of the direction of transfer." (p.518) Hierarchical structure in HCAs is *ontological*, that is, the CA substrate is itself hierarchically structured. This contrasts strongly with, for example, UCAs in which it is (emergent or designed) *phenomena* which are hierarchically structured via embedding of virtual machines in a substrate which is itself ontically non-hierarchical in structure. (It is logically possible that HCAs might also support an embedded virtual machine hierarchy. In this case, such a HCA would be regarded as *doubly-hierarchical* since both its ontology and phenomenology would be hierarchically structured.) The difference between a HCA and a UCA supporting an embedded machine hierarchy is shown in Fig 5.4:

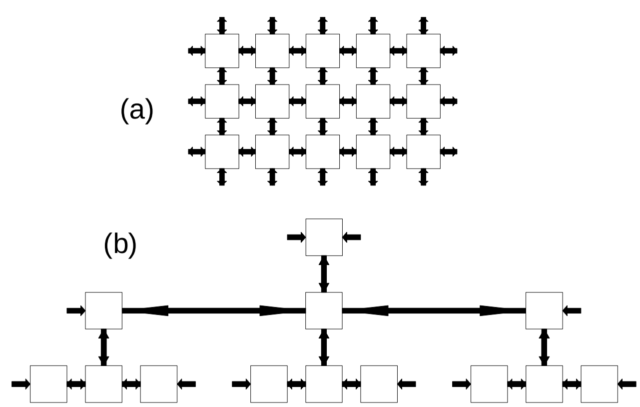


Fig 5.4 (a) standard two-dimensional CA; (b) two-dimensional HCA.

[30] This is consistent with Cairns-Smith's (1985) notion of genetic takeover in the context of the origin of life problem.

5.5.2. Why CEA *is* Possible

CEA is *possible* because computationalism or computational-functionalism (chapter 2) supports the dualistic decoupling of emergent(s) from substrate, phenomenon from noumenon, appearance from grounding reality and hence, allows for multiple realization (chapter 4); computationalism is *possible*, that is, not *a priori* objectionable on logical grounds, since it is a `relatively adequate' (albeit eclectic) world hypothesis (chapters 1 and 2). The link between emergence (chapter 3) and the hardware-software dualism supported by computationalism (chapter 2) is described by Hillis (1988) in the context of the philosophy of mind as follows, viz. "it seems likely that symbolic thought can be fruitfully studied and perhaps even created without worrying about the details of the emergent system that supports it." (pp.179-180)

> Although the detailed rules of interaction are very different from the interactions of real molecules [in a fluid], *the emergent phenomena are the same*. The emergent phenomena can be created without understanding the details of the forces between the molecules or the equations that describe the flow of the fluid [emphasis added]. (p.187)

However, it is important to appreciate at the outset that the *range* of possible candidate substrates will, in fact, be highly constrained, the multiple-realizability thesis notwithstanding. This is because any candidate must demonstrate (functional and behavioural) *sufficiency* with respect to its role as a substrate for the computational emergence of artificial analogues of natural phenomena such as matter, life and mind.

5.5.3. Towards CEA: The Game of Life (GOL)

Conway's Game of Life (GOL), which is capable of supporting the computational emergence (chapter 3) of isomorphic analogues (functional, behavioural) of physical (material), biological (vital) and psychological (mental) phenomena, is, *a priori*, a viable candidate for the realization of CEA. While there are a number of requirements which must be met in order that the GOL support CEA, for example, provision of a necessary minimum space and time for phenomenal production, precision in initial organizational configuration (since, as will be shown in section 5.5.9, the GOL is highly non-robust) etc, nonetheless, the GOL does, *in principle*, provide what is, *a priori*, a viable ontological substrate for the implementation of a unified framework of CEA. Detailed accounts of the phenomena which can be generated by the GOL are presented in (Berlekamp,82), (Gardner,83) and (Poundstone,86). Table 5.4 is a non-exhaustive list[31] of some basic and derivative[32] types of GOL structure:

---

31 For example, the breeder (Poundstone,86), one of the most complex GOL structures after embedded virtual computers (Berlekamp,82), has not been included.

32 Glider-guns and other higher-order structures are considered *derivative* (or complex) since they can be constructed from a combination of suitably positioned translating oscillators and static structures

Phenomenal Class	*Description*
Blankers	unstable structures
Blocks, Beehives etc	static structures
Blinkers, Flip-Flops etc	oscillators
Gliders, Spaceships etc	translating-oscillators
Glider-guns	translating-oscillator generators
Puffer-trains	translating translating-oscillator generators

Table 5.4 Basic and derivative types of GOL structure.

The rich phenomenology of the GOL CA makes possible the establishment of isomorphisms (functional, behavioural) between various elements in Alexander's emergentist metaphysics and those which of necessity must appear in a CA-computationalism supporting CEA. A non-exhaustive list of these correspondences is shown in table 5.5:

	Alexanderian Metaphysics	*CA-Computationalism (GOL-based)*
Ontological Substrate	a single infinite plurality of point-instants (events)	a single infinite lattice of cells each containing a finite state machine
Matter	moving patterns of events	translating oscillators or gliders
Life	moving patterns of matter	self-reproducing structures or breeders (Poundstone,86)
Mind	moving patterns of life	computation-universal structures or embedded virtual computers (Berlekamp,82)

Table 5.5 Isomorphisms between Alexander's metaphysics and CA-computationalism.

At this point in the discussion, it is worthwhile examining the GOL realization of CEA in slightly more detail. This will provide a clearer understanding of *how* computationalism attempts to subsume earlier materialist conceptions of naturality.

5.5.3.1. GOL-Matter

Berlekamp et al. (1982) maintain that if CA-computationalism is correct and the universe is a computer then

> what we call motion may be only *simulated* motion. A moving particle in the ultimate microlevel may

(Berlekamp,82).

> be essentially the same as one of our gliders, *appearing* to move on the macrolevel, whereas actually there is only an alteration of states of basic space-time cells in obedience to transition rules that have yet to be discovered [emphasis added]. (p.849)

This position is consistent with the `primal-stuff' interpretation of Space-Time proposed by Alexander and described in section 5.4.2. Additionally, the GOL can be extended so as to incorporate various empirical features associated with the physical world; for example, Bays (1987a, 1987b, 1990, 1991, 1992) presents a number of three-dimensional versions of the GOL based on cubic and spherical lattice geometries. Importantly, it has been shown that a cubic variant can be structured so as to contain an infinite number of parallel two-dimensional universes, each of which allows for the evolution of Conway GOL objects, thereby demonstrating the GOL's capacity for the embedding of computers and the construction of virtual machine hierarchies (chapters 2 and 4).

5.5.3.2. GOL-Life

The possibility of self-reproduction is entailed by the computation universality of the GOL as anticipated by von Neumann (section 5.5.7). As Berlekamp et al. (1982) state,

> Eaters [that is, static glider-consumers] and guns can be made by crashing suitable fleets of gliders, so it's possible to build a computer simply by crashing some enormously large initial pattern of gliders. Moreover, we can design a computer whose sole aim in Life is to throw just such a pattern of gliders into the air. *In this way one computer can give birth to another*, which can, if we like, be an exact copy of the first [emphasis added]. (p.848)

Moreover, a CA-based variant of evolution via natural selection is also logically possible, viz.

> among finite Life patterns there is a very small proportion behaving like self-replicating animals. Moreover, it is presumably possible to design such patterns which will survive inside the typical Life environment (a sort of primordial broth made of blocks, blinkers, gliders, ...). It might for instance do this by shooting out masses of gliders to detect nearby objects and then take appropriate action to eliminate them. So one of these `animals' could be more or less adjusted to its environment than another. If both were self-replicating and shared a common territory, presumably more copies of the better adapted one would survive and replicate. (p.848)

> [Hence,] it's probable, given a large enough Life space, initially in a random state, that after a long time, *intelligent, self-reproducing animals* will emerge and populate some parts of the space [emphasis added]. (p.849)

The latter statement assumes that intelligence is *continuous* with life (chapter 4) and leads directly to the GOL interpretation of mind.

5.5.3.3. GOL-Mind

The GOL supports universal computation (chapter 2). Furthermore, using logic gates (NOT, AND, OR) constructed from suitably positioned glider guns and other basic structures, it is possible to ground or embed a virtual computer, that is a CA equivalent of a universal Turing machine, in the GOL substrate. A constructive proof is presented in (Berlekamp,82). Computation-universality is *directionally* important for (at least) two reasons: (1) `upwardly' because it has been correlated with intelligence in computationalist theories of the mind and is readily extended to cope with the phenomenon of intentionality if an objectivist-behaviourist interpretation of the latter is adopted (chapter 4). The phenomenon of consciousness (subjectivity, first-personhood etc) is explained in emergentist terms: Following Alexander, CA-computationalists maintain that consciousness is an emergent phenomenon associated with a specific organization of its causal substrate, viz. the brain; (2) `downwardly' because computation-universality supports *simulability* (chapter 4) and, thereby, the *simulation* (or realization) of other phenomena characteristically associated with life such as autonomy, growth, metabolism and evolution.

5.5.4. Problems with the GOL

While it might appear that the GOL is sufficient (functionally, behaviourally) as a substrate for CEA, close examination reveals this position as both naive and incorrect. One of the main problems with the GOL with respect to CEA is that it is informationally-*dissipative*, that is, *irreversible*; however, as Gutowitz (1990) states, "fundamental physical laws are (microscopically) time-reversal invariant, hence a cellular automaton which is to model such physics should be time-reversal invariant as well." (p.viii) Thus, the GOL is not a suitable substrate for realizing microscopic physics. In addition, GOL causation does not support quantum non-locality and relativistic effects (chapter 4) are not readily incorporated due to the scale invariance of structures in the CA. While it is true that the GOL can support the virtual machine embedding of, for example, Newtonian (classical or reversible) systems, this is not an *efficient*[33] approach. For this reason, it is necessary to extend the basic CA model to incorporate the requisite features as described below.

5.5.5. Reversible and Irreversible CAs

In the previous section, notions of reversibility and irreversibility were introduced in the context of CAs. Reversibility is a universal characteristic of physical law for phenomena above the quantum level and is a *necessary* condition for the Second Law of Thermodynamics to hold. As Wolfram (1984b) states, "the laws of thermodynamics give a general description of the overall behaviour of systems governed by microscopically non-dissipative (reversible) laws." (p.viii) In locally-interacting systems which have a

---

33 `Efficient' in the sense of parsimonious. Since there are no *a priori* reasons necessitating that the state-transition function of the CA substrate implementing CEA be irreversible, therefore, in accordance with Ockham's Razor, a reversible state-transition function will be assumed. This entails a *simplification* of the relationship between substrate and first level emergent, viz. matter.

finite amount of information per site (for example, standard CAs), reversibility is *equivalent* to the Second Law. Reversibility also provides a *sufficient* condition for the conservation of various physical quantities such as energy, momentum etc (Toffoli,87). CAs are *forwards-deterministic* systems, that is, for every global CA state, the CA rule specifies a unique successor state[34]. *Reversible* (or invertible) CA are both forwards *and* backwards deterministic; hence, a unique predecessor state exists for all CA states except Garden of Eden configurations (chapter 2). The rule allowing the trajectory (or global state-sequence) to be traversed in reverse is known as the *inverse* rule with respect to the original forward rule. (A CA is called *time-reversal invariant* if the forward and inverse rules are identical under a suitable transformation of the final state generated by the forward rule into the initial state of the the reverse rule.) Margolus (1984) defines *reversible cellular automata* (or RCAs) as

> computer-models that embody discrete analogues of the classical-physics notions of space, time, locality, and microscopic reversibility. (p.81)

Toffoli and Margolus (1990) hold that bijectivity[35]constitutes a *necessary* condition for reversibility in dynamical systems where reversibility in CAs is defined as follows:

> A cellular automaton is invertible [or reversible] if its global map is invertible, i.e. if every configuration - which, by definition, has exactly one successor - also has exactly one predecessor. In the context of dynamic systems, invertibility coincides with what the physicists call `microscopic reversibility'. (p.231)

However, in CAs injectivity[36] is equivalent to invertibility (reversibility) since CAs are deterministic (chapter 2) and hence, all states are reachable in the *forwards* direction of spatio-temporal evolution[37]. Thus, CAs are a special case of dynamical systems in which bijectivity *reduces* to injectivity.

Reversibility of CAs has been widely investigated in the literature. Toffoli and Margolus (1989) present a general introduction to the concept of invertibility and detailed presentations of the concept have been made in (Fredkin,82,90,96), (Margolus,84,93) and (Toffoli,77,78,80,89,90,94a,94b). However, before examining how invertibility can be implemented in CAs, it is worthwhile considering the following remark due to Capecci (1979), viz.

---

[34] This corresponds to a surjective mapping (see Glossary).

[35] See Glossary.

[36] See Glossary.

[37] This includes Garden of Eden states (chapter 2) which are axiomatically reachable.

> it is impossible to establish a reversible relation of cause and effect except when we are dealing with entities which are *eternal* and which are achieved by *abstraction* [emphasis added]. (p.38)

If this statement is correct - and to maintain that this is the case necessitates holding with Prigogine and Stengers (1984) that the "arrow of time" (directionality) associated with thermodynamic processes (chapter 4) is ontical as opposed to merely epistemic - then it is possible that both Newtonian mechanics and RCAs are abstract epistemological constructs (that is, artifacts), products of a `cutting' of the world that ignores its intrinsic temporality or *historicity* (chapter 6). In the context of the present discussion, this view will be regarded as in conflict with the reversibility and determinism of both the classical and quantum *formalizations* of microscopic phenomena (chapter 4) and hence, will not be considered further.

5.5.5.1. Implementing Reversibility

If a CA *is* invertible, then its inverse is a cellular automaton and a local FSM state-transition map for the inverse CA can be defined; however, the invertibility of a n-dimensional CA is undecidable (chapter 2) for $n \geq 1$ (Toffoli,90). For this reason, *constructive* or synthetic (as opposed to *deductive* or analytic) approaches have been adopted in the context of RCA implementation. Toffoli and Margolus (1987) describe a number of possible schemes for implementing reversibility in CA, for example, alternating sublattices, the guarded-context technique etc. However, possibly the simplest and most general approach is the *second-order technique* suggested by Fredkin which involves specifying the forward-time dynamics of a CA as

$$s^{t+1} = \tau(s^t) - s^{t-1} \qquad (5.1)$$

where s^t is the state of the CA at time t and τ is the forward state-transition rule. (5.1) guarantees the reversibility of the dynamics for an arbitrary τ. The reverse-time dynamics are given in (5.2) and are determined by solving (5.1) for s^{t-1}. Time-reversal invariance is supported by applying $\tau \equiv \tau^{-1}$ to suitably defined time-reversed states.

$$s^{t-1} = \tau(s^t) - s^{t+1} \qquad (5.2)$$

Importantly, the global time evolution generated by (5.1) is not guaranteed to be invertible unless suitable boundary conditions are chosen, such as no boundary (that is, an infinite or periodic space) or `fixed' boundaries (cell values on the boundary are not allowed to change with time) (Margolus,84). The above scheme is important since it shows that in spite of the rarity of reversible CAs (or RCAs), the ones that are constructable are *at least* as many as the corresponding irreversible or nonreversible CAs (or NRCAs) (Toffoli,90).

Yet another approach to RCA implementation involves the notion of partitioning. In addition to providing a means by which to implement reversibility in CA, the partitioning technique is significant in connection with the debate over the relative autonomy of levels in a phenomenal hierarchy. A *partitioned cellular automaton* (PCA) is a variant of standard CA in which

1. the lattice is partitioned into a collection of finite, disjoint and uniformly arranged *blocks*;

2. a *block-transition rule* is defined and uniformly applied to each block in the lattice. Blocks are not allowed to overlap and information is not exchanged between adjacent blocks;

3. the *block-partitioning* is changed at each iteration such that there is overlap between the blocks used at consequtive iterations. This ensures casual connectivity (i.e. dependency) of cells in the lattice.

Hence, a PCA is essentially a standard CA with `coarse' *granularity*. (Standard CAs are, by contrast, `fine' grain since the block size collapses to that of a lattice primitive, that is, a unit cell). Space-time uniformity is achieved by the cyclicity of the partitioning strategy. It is important to note that PCAs can be implemented in standard CAs using more states and neighbours (Toffoli,87). The inverse rule is generated by applying the forward rule in reverse, that is, $b^t \rightarrow b^{t+1}$ becomes $b^t \leftarrow b^{t+1}$ or rather $b^t \rightarrow b^{t-1}$, where b^t denotes a partitioning block of states at time *t*. However, there must be a one-one correspondence between `old' and `new' block states such that the rule consequent is a permutation of the rule antecedent. The distinction between NRCAs and PCAs is shown in Fig 5.5:

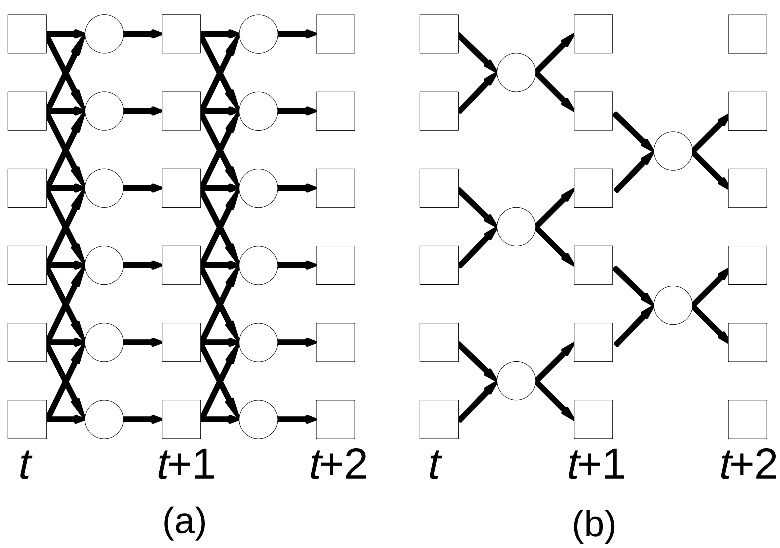


Fig 5.5 (a) NRCA; (b) PCA.
(Squares denote cells, circles denote operators.)

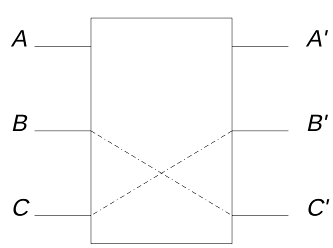
5.5.5.2. Billiard-Ball Mechanics (BBM)

Features associated with both schemes are integrated in the `billiard-ball' model of CA-computation or BBM-CA (Margolus,84) which makes use of the notion of *conservative* *logic* (Fredkin,82). This implies conservation in the output of the number of 0's and 1's that are present in the input to a *gate*, viz. a primitive RCA element composed from a block of lattice cells. According to Toffoli (1980),

> in conservative logic, all data processing is ultimately reduced to *conditional routing* of signals. Roughly speaking, signals are treated as unalterable objects that can be moved around in the course of a computation but never created or destroyed. (p.640)

Hence, the consistency of conservative logic with the microphysical basis underlying the laws of thermodynamics (chapter 4). Furthermore, as Landauer (1967) states, conservative logic is consistent with the non-production of information. This point is extremely important in the context of emergence and self-organization which are usually interpreted as information-generating (or *negentropic*) processes. (However, it is important to recognize that RCAs cannot support self-organization *by definition* since they are non-surjective[38] automata (Wolfram,84b).)

Conservative logic is essentially a type of logic which is consistent with the physics of microscopically reversible or classical (Newtonian) systems. The *billiard-ball model* (BBM) of computation (Fredkin,82) is a classical mechanical system obeying a continuous dynamics; more precisely, a system of `hard' (momentum-conserving) spheres and `mirrors' (or reflecting surfaces). By suitably restricting the initial conditions the system can assume and by only looking at the system at regularly spaced time intervals, the BBM can be made to perform a digital process. According to Margolus (1984), "the key insight behind the BBM is this: every place where a collision of finite-diameter hard spheres might occur can be viewed as a boolean logic gate." (p.86) The *Fredkin gate* (Fredkin,82) shown in Fig 5.6 implements a BBM logic gate. This gate, which is its own inverse, is a reversible (invertible) universal logic element; thus, any invertible logic function can be constructed out of Fredkin gates.



A	*B*	*C*	*A′*	*B′*	*C′*
0	0	0	0	0	0
0	0	1	0	1	0
0	1	0	0	0	1
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	0
1	1	1	1	1	1

Fig 5.6 The Fredkin (Reversible Logic) Gate.

[38] Non-surjective here implies injectivity and hence, non-existence of a basic of attraction field.

For example, by presetting line *B* to `0' and discarding *A′* and *B′* on the output, the AND function can be computed. Furthermore, as Wolfram (1984b) states,

> cellular automaton rules may be combined by composition. The set of cellular automaton rules is closed under composition, although composition increases the number of sites in the neighbourhood. (p.3)

Hence, by cascading a NOT gate (inverter) and an AND gate, a NAND gate can be constructed (Toffoli,80). This is important since NAND gates (like NOR gates) are *universal* logic elements which can be used as the basic (or `atomic') primitives from which to construct computers (UTMs). However, as Toffoli and Margolus (1990) state,

> in an *invertible* cellular automaton the gates [or local FSM state-transition maps] will have to be invertible; because of this constraint, a complete, self-contained logic design will have to explicitly provide, besides circuitry for the desired logic functions, additional circuitry for functions (analogous to energy supply and heat removal in ordinary computers) concerned with entropy balance. (p.234)

By taking a long time to perform the computation, a physical realization of a reversible logic gate can expend arbitrarily little energy. Consequently, Bennett and Landauer (1985) maintain that "there is thus no minimum amount of energy that must be expended in order to perform any given computation." (p.41) While this result makes possible more efficient designs of computers with respect to energy constraints, it does not entail that energy expenditure can be reduced to zero; physical realizations of computers have moving parts and hence, generate friction which leads to the dissipation of energy in the form of heat. It would appear, therefore, that only abstract (`Platonic') RCAs can attain zero energy expenditure. Toffoli (1980) admits this in stating that "it appears possible to design circuits whose internal power dissipation under *ideal* physical circumstances is zero [emphasis added]." (p.633) Furthermore, if Prigogine and Stengers (1984) are correct and microscopic fluctuations constitute an irreducible feature of the physical world then classical thermodynamics and the notion of reversibility are idealizations which apply to certain stable macroscopic structures such as those for which classical mechanics was developed as a descriptive device. Bennett and Landauer (1985) point out additional problems with systems based on reversible logic:

> For example, do elementary logic operations require some minimum amount of time ? What is the smallest possible gadgetry that could accomplish such operations ? Because scales of size and time are connected by the velocity of light, it is likely that these two questions have related answers. We may not be able to find these answers, however, until it is determined whether or not there is some ultimate graininess in the universal scales of time and length. (p.45)
>
> At the other extreme, how large can we make a computer memory ? How many particles in the universe can we bring and keep together for that purpose ? The maximum possible size of a computer memory limits the precision with which we can calculate. (p.45)

If the universe is a finite UCA, then it only *approximates* the computational capability of a UTM (chapter 2). (A finite UCA is equivalent to a UTM with a large but finite tape.) Thus, there will be numerical entities (such as pi) whose expansion can only be *computed* to finite precision. This point is important: If the universe is infinite, the complete expansion of pi can be generated. However, in a finite universe it cannot, which leads to an interesting question, viz. is pi *really* infinite or is it equal merely to its computable expansion, an expansion whose length is at least partially determined by the physical constraints on computation imposed by the universe ? Such questions lead ultimately into metaphysical territory with Platonists such as Penrose (1994) subscribing to the former position and instrumentalists (finitist-constructivists) such as Cariani (1989) subscribing to the latter. Bennett and Landauer (1985) advance a variant of the latter view supporting the postulated closure between physics, mathematics and computation described in chapter 4:

> The inevitable deterioration processes that occur in real computers pose another, perhaps related, question: can deterioration, at least in principle, be reduced to any desired degree, or does it impose a limit on the maximum length of time we shall be able to devote to any one calculation ? That is, are there certain calculations that cannot be completed before the computers' hardware decays into uselessness ? Such questions really concern limitations on the physical execution of mathematical operations. Physical laws, on which these answers must ultimately be based, are themselves expressed in terms of such mathematical operations. *Thus we are asking about the ultimate form in which the laws of physics can be applied given the constraints imposed by physics that the laws themselves describe* [emphasis added]. (pp.45-46)

5.5.5.3. BBM-CA

A RCA implementing the BBM (a BBM-CA) is readily constructed since RCAs model space as being uniformly filled with uniformly-connected gates (Margolus,84). Toffoli (1980) maintains that in *combinational* (or causally non-iterative) networks, structural reversibility and functional invertibility coincide whereas in *sequential* (or causally iterative) networks, the network is said to be reversible if its combinational part is reversible. On this basis, Toffoli and Margolus (1990) conjecture that *all* RCAs are structurally invertible, that is, can be (isomorphically) expressed in spacetime as a uniform composition of finite invertible logic primitives. As an example of the BBM-CA, consider the following two-dimensional binary RCA: The (Cartesian) lattice is divided into 2×2 blocks of cells and each block is treated as a conservative logic gate with four inputs (current state) and four outputs (next state). The local update rule shown in Fig 5.7(a) is applied to the 2×2 blocks using alternately the solid and dotted blocking regimes. The application of this rule in the modeling of particle collisions[39] is shown in Fig 5.7(b).

39 It should be noted that this technique introduces a 1-block delay in the result as shown.

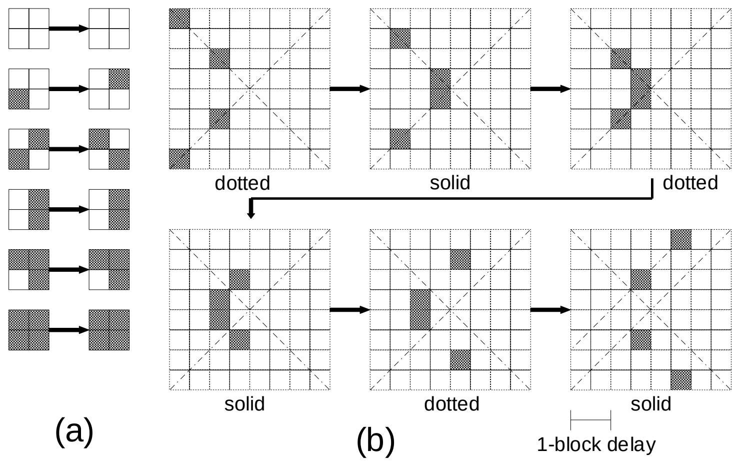


Fig 5.7(a) BBM-CA rule; (b) particle collision.

In the following sections, CA realizations of the computational theories of matter, life and mind discussed in chapter 4 will be presented. Implementation of the artificial analogue of each natural phenomenon is not described in detailed, nor are the artificialities integrated under a single CA. It is assumed that since CA-computationalism in either its RCA or NRCA form can realize each phenomenon individually, a unified framework of CEA is logically possible. The grounds for this assumption are examined in section 5.5.9.

5.5.6. CA-Matter

Matzke (1996) defines *information* as "just another kind of energy or matter (or visa versa)" and holds that "information laws are toplogical constraints that precede physical laws." (p.223) This is consistent with Wheeler's notion of a pregeometric foundation for physics, briefly mentioned in the context of the computational theory of matter or CTMa (chapter 4). The CTMa can be implemented under RCA-computationalism using the digital mechanics framework introduced in chapter 2 and described in more detail below. As a precursor to the presentation of this scheme, the nature of the link between CAs and physics will be examined.

Vichniac (1984) distinguishes three ways in which CAs may be used to simulate physics, viz. "(1) CAs as computational tools; (2) CAs as fully discrete dynamical systems. In this approach, cellular automata are relevant to physics only insofar as dynamical systems are relevant to physics; (3) CAs as original models for actual physical phenomena, possibly competing with existing continuum models." He goes on to state that

> many of these CA universes are inhabited by beings that are most often seen in the theoretical physicists menagerie, such as symmetries and conservation laws, a conspicuous arrow of time in

> reversible microscopic dynamics, order parameters and non-ergodicity, non-separability, causality and light cone, relaxation to chaos through period doublings, and, most instructively, the appearance of complex phenomena and large scale correlations resulting merely from a very simple short-range interaction. (p.97)

Toffoli and Margolus (1987) describe the application of CAs in the modeling of various physical phenomena including diffusion, fluid dynamics, Ising system dynamics and, implicit in the discussions of section 5.5.5, classical and thermodynamic processes. Ablowitz et al. (1991) describe a class of stable time-reversible multistate CAs possessing a large array of coherent particle-like structures governed by rich particle-interactions including particle production. However, there are a number of phenomena that have not been successfully incorporated into RCAs. For example, according to Margolus (1993), relativity is not supported since "there is no real notion [in a CA universe] of the `same' composite system in various states of motion." (p.3) Although CA models of quantum-mechanical (or QM) phenomena such as wave-particle complementarity (Gernert,86) and analogues of Heisenberg uncertainty (Ingerson,84) have been described, it is generally held that there is a limit to the modeling capabilities of CAs with respect to QM. According to Davies (1992),

> Heisenberg's uncertainty principle puts paid to the notion that the present determines the future exactly. Of course, this supposes that quantum uncertainty is genuinely intrinsic to nature, and not merely the result of some hidden level of deterministic activity. In recent years a number of key experiments have been performed to test this point, and they have confirmed that uncertainty is indeed inherent in quantum systems. The universe really is indeterministic at its most basic level. (pp.30-31)

The same point is made in a computational context by Wolfram (1983a), viz.

> the Church-Turing thesis states that no system may have computational capabilities greater than those of universal computers. The thesis is supported by the proven equivalence of computational models such as Turing machines, string-manipulation systems, idealized neural networks, digital computers, and cellular automata. While mathematical systems with computational power beyond that of universal computers may be imagined, it seems likely that no such systems could be built with physical components. This conjecture could in principle be proved by showing that all physical systems could be simulated by a universal computer. *The main obstruction to such a proof involves quantum mechanics* [emphasis added]. (p.17)

Furthermore, Feynman (1982) has *proved* mathematically that classical computers can never simulate non-locality. Margolus (1993) presents the following list of characteristics incorporated into CAs from computation theory and physics as well as a number of features which are absent in CA-Matter (Table 5.6).

Computation	*Physics*	*Missing*
digital exact universal	space time locality of interaction finite info^n. propagation speed 3-dimensional interconnectivity conservation laws microscopic reversibility finite entropy for finite system entropy proportional to volume	relativity quantum mechanics

Table 5.6 Computational and physical characteristics of CA-Matter.

5.5.6.1. Digital Mechanics (DM) Revisited

Digital Mechanics or DM (Fredkin,90) refers to CAs which have been programmed to act like physical systems. Fredkin assumes the finite nature thesis (chapter 2), the main implication of which is that a finite volume of space-time can be exactly represented by a finite number of bits; thus, physics can be reduced to discrete computation. He maintains that "DM like systems may be useful for creating approximate models of continuous phenomena, but if finite nature is true, then DM can be an *exact* model [emphasis added]" (p.255) in which case simulation gives way to emulation (chapter 4). DM is described as follows, viz.

> to be complete, everything in our universe (the exact present state) would be a consequence of the rule, the size and shape of the cellular space, the boundary conditions (which can be eliminated if the space wraps; e.g., if cell coordinates are treated modulo some number), the initial conditions and the time (which would be an integer, the number of time steps from the initial conditions to the present. (p.258)

According to Fredkin, the implications of this thesis are as follows:

> finite nature would mean that our world is an informational process - there must be bits that represent things and processes that make the bits do what we perceive of as the laws of physics. This is true because the concept of computational universality guarantees that if what is at the bottom is finite, then it can be exactly modelled by any universal machine. (p.259)

The following arguments in support of finite nature, viz.

> as an informational scientist it is not that we see DM as a possible model of physics, rather it is that we see no way to model physics without the incorporation of much of what is in DM. In other words, to a programmer who believes in finite nature, physics cannot be imagined in microscopic detail without its having most of the characteristics of DM; unless one resorts to magic. A programmer who does not believe in finite nature, knows that he can only model physics on a computer by making the same kinds of assumptions that DM makes. (p.259)

DM and computation-universal RCAs (or RUCAs) are closely connected. According to Fredkin,

> for each RUCA, there are one or more .. informational processes [which can supervene on the RUCA substrate]. DM is simply an informational process that runs on a RUCA and that in one way or another should be able to model nature. (p.259)

Fredkin presents the following list of correspondences between DM (grounded in some supporting RUCA) and physics (Table 5.7):

Physics	*DM*
digit-transition	D
length	L, the inter-cell distance
time	T, one CA clock cycle
energy	D/T
momentum	D/L
mass	DT/L^2
angular momentum	D
action	
charge (+ or -)	space-time parity of D, even or odd
charge quantization	stable D orbits in 3-space
color	structure orientation:N-S, E-W, U-D
2-state system (spin)	actually, measuring one bit!
conservation laws	conservation of information
isotropy	asymptotic isotropy
continuity	discreteness
infinitesimals	the digit, units of length and time
infinities	large but finite
special relativity	asymptotic special relativity
general relativity	consequence of the DM process
measurable acceleration	measurable velocity
measurable rotation	measurable angular orientation
group theory properties	consequence of RUCA symmetries
particle masses	stabkle structures in the RUCA
too many parameters	the rule, and the initial conditions
why is there anything ?	answer: unknowable determinism
compex amplitudes	2-phase clock, time dimensions depth 2
spin 1/2	smallest D orbit
isotopic spin	projection of D orbits that represent charge
form of the photon	each particle is a digital machine where its
form of the electron	spin, momentum energy, charge, color etc.
form of the quark	is represented by information or by a
form of the gluon	particular information process
form of the neutrino	

Table 5.7 Correspondences between Physics and DM (Fredkin,90).

Cursory examination of this table shows DM to be capable of modeling (or realizing on the "strong" computationalist view) both relativistic and quantum mechanical phenomena. However, these issues warrant further consideration. With respect to relativity, Toffoli (1989) maintains that

> quantitative features of special relativity and at least qualitative features of general relativity emerge quite naturally as epiphenomena of very simple computing networks [flat, uniform lattice gas CAs]. Thus, relativity appears to be of the right form to be an emergent property, whether or not *that* is the way it comes about in physics. (p.317)

However, Fredkin (1990) states that

> we now know of reasonable approaches that allow the physics of DM to be Lorentz invariant [that is, relativistically correct] even though the process runs on a Cartesian lattice. The lattice remains Cartesian, but measurements made *from within* the DM model by sending simulated photons back and forth to make measurements can be expected to give relativistically correct results [emphasis added]. (p.258)

Furthermore,

> a DM system can be asymptotically isotropic. Nearly all trace of the preferred space-time coordinate system, its anisotropy, its absolute reference frame and its absolute lengths and times can be totally washed out so as to become relativistically correct as the scale of events moves away from the most microscopic. (p.262)

Thus, by adopting an *endophysical* (or intra-systemic) as opposed to *exophysical* (or extra-systemic) perspective, that is, by embedding the observer *within* the DM-RUCA universe, Fredkin maintains that relativistic effects will be observed; DM-RUCAs are absolute (Newtonian, classical) only from the privileged perspective of an external observer who can assume a "Gods-Eye" view of the DM-RUCA universe. This is *possible* for human beings with respect to DM-RUCAs since human beings are the artificers of these virtual worlds and hence, stand outside of (or beyond) them.

With respect to QM, Fredkin (1990) maintains that "DM may be capable of using mechanistic, deterministic, and local rules as a substrate and yet produce behaviour that obeys the laws of QM." (pp.259-260) How is this possible ? Fredkin claims that the conflict between classical and quantum mechanics is merely apparent. By adopting a variant of the `hidden variables' interpretation of QM (chapter 4) and utilizing a distinction between light cones and information cones, Fredkin is able to explain quantum non-locality under DM:

> In DM, there is an information cone that is loosely equivalent to our current idea of the light cone of ordinary physics. However, the information cone does not have the intuitive kind of local causality we sometimes attribute to the light cone. *In DM it is simply incorrect to say that only events within the information cone of the past can influence an event in the present*. This is surprising, but really has to do with the nature of DM. The state of a particular cell is absolutely determined by the state of its immediate neighbours in space-time [emphasis added]. (p.264)

> ... in every reversible CA, the space-time neighbours of a cell always include both the past and the future .. this means that there are reasonable DM models where every cell is a deterministic function of cells through a region of space, the information cone, that includes parts of space *not in the obvious light cone*. Thus the language we use to describe situations where the concept of a light cone is used, may make little sense in DM [emphasis added]. (p.264)

Using the second-order construction technique (section 5.5.5), the various components in the *information* cone for a RUCA cell can be defined as shown in Table 5.8:

Computing the	*Formalism*
future	$C_{x,t+1} = F_1(C_{x-1,t}, C_{x+1,t}, C_{x,t-1}, C_{x,t})$
present	$C_{x,t} = F_2(C_{x-1,t}, C_{x+1,t}, C_{x,t-1}, C_{x,t+1})$
past	$C_{x,t-1} = F_3(C_{x-1,t}, C_{x+1,t}, C_{x,t+1}, C_{x,t})$
right (of *light* cone)	$C_{x+1,t} = F_4(C_{x-1,t}, C_{x,t}, C_{x,t-1}, C_{x,t+1})$
left (of *light* cone)	$C_{x-1,t} = F_5(C_{x+1,t}, C_{x,t}, C_{x,t-1}, C_{x,t+1})$

Table 5.8 Computing the information cone in a RUCA cell.

Consequently,

> the present is no more a consequence of the past than it is a consequence of the future. A location is no more the immediate consequence of its local neighbourhood than it is a consequence of very distant points in space-time. Consider a given event: if we have partially determined certain states in its past [via the imposition of *local* initial conditions] and arrange that in its future certain other states will be determined, those states that will be determined in the future will have an effect on the event! (pp.264-265)

Furthermore,

> the past can only be totally determined for an event if the total state of all of space-time is determined. (p.265)

Fredkin maintains that "this very property of DM can allow for mechanistic models of the mysterious events of QM" (p.265) such as non-locality since "in DM, everything is, in some sense, involved in a computation with everything else." (p.267) Thus, DM implies that "QM can be modelled by Newtonian mechanics; but not in a way that would easily have occurred either to Newton or to the average quantum mechanic." (p.269)

5.5.7. CA-Life

The computational theory of life (CTL) was described in chapter 4 and a number of properties such as growth, metabolism, autonomy, self-repair, self-reproduction and evolution were identified as characteristic of biological systems. If the CA substrate supports CEA (in this case the emergence of life) and if the substrate is computation-universal (that is, if the CA is a UCA) then computational analogues of each of these properties can be constructed. Put simply, if life is algorithmic, CA-Life is logically possible. Various CA models of biological phenomena have been reported in the literature: For example, Burks and Farmer (1984) have investigated the issues associated with modeling the evolution and functionality of DNA sequences using CAs. This is possible because "the DNA molecule can be viewed as a one-dimensional lattice, with four states per lattice site, that is capable of producing copies of itself." (p.158); Ermentrout and Edelstein-Keshet (1993) describe various uses of CAs in the modeling of biological phenomena. However, in what follows, only three properties associated with CA-Life, viz. self-reproduction, evolution and autonomy will be considered for purposes of illustration.

5.5.7.1. Self-Reproduction

Von Neumann[40] (1966) was interested in determining what kind of logical organization is *sufficient* for an automaton to be capable of self-reproduction[41]. Langton (1984) describes his approach as follows:

> If self-reproduction *is* being carried out by a (highly complex) biochemical machine, then that machine's behaviour is describable as a logical sequence of steps, i.e. as an algorithm. Now, if an algorithm can be performed by any machine at all, then there is a Turing machine which can perform the same algorithm. For this reason von Neumann set out to demonstrate the existence of a Turing machine which could affect its own reproduction. If such a Turing machine exists, it is entirely plausible that the processes by which living organisms reproduce themselves, and by implication, other processes on which life itself is based, are algorithmically describable and, therefore, that life itself is achievable by machines (p.135)

Von Neumann presented a construction- and computation-universal self-reproducing structure using a two-dimensional CA in which each cell could assume one of twenty nine states and in which the cell neighbourhood comprised five cells. Later work by Codd (1968) demonstrated that it was possible to reduce the number of cell states (to eight) while retaining universality. Langton goes on to state that

> [von Neumann's] Turing machine is suitably modified so that, as output, it can `construct' in the array [that is, the CA lattice] any configuration which can be described on its input tape. Such a machine is called a *universal constructor*. His machine will construct any machine described on its input tape and, in addition, will also construct a copy of the input tape and attach it to the machine it has

40 Cariani (1989) has contested the computationalist claim that von Neumann himself subscribed to a computationalist interpretation of self-reproduction and evolution. Cariani maintains that according to von Neumann, the genotype-phenotype distinction in biological systems corresponds to an analog-digital (or in Cariani's terms, a nonsymbolic-symbolic) distinction. If computation is understood as implying *digital* (or discrete) symbol manipulation - as is held in this thesis (chapter 2) - then Cariani's interpretation of von Neumann's position is clearly correct.

41 It is important to note that there are two levels of automata: (1) the embedd*ing* CA substrate and (2) the embedd*ed* self-reproducing automaton.

> constructed. Now, self-reproduction follows as the special case where the machine described on the tape is the universal constructor itself. The result of the construction process is a copy of the universal constructor together with an input tape which contains its own description, which can then go on to construct a copy of itself, together with a copy of its own description, and so on indefinitely. (pp.135-136)

The information on the `tape' is used in two ways: (1) as instructions (for construction) which are to be *interpreted* (corresponding to the genotypic process of *translation* in biological organisms) and (2) as *uninterpreted* data which is copied directly and attached to the newly constructed automaton (corresponding to the genotypic process of RNA *transcription*). Although computation- and construction-universality is a *sufficient* condition for self-reproduction, it cannot be a *necessary* condition since this would entail viewing natural biological systems as non-self-reproductive, for as Langton states, "none of these have been shown to be capable of universal construction." (p.137) This leads Langton (1987) to distinguish between *trivial* and *non-trivial* reproduction as follows: the former involves simple (passive) replication whereas the latter involves complex (active) self-reproduction mediated by a representation or description of the reproducing/reproduced entity. On this view, non-trivial non-universal self-reproduction is possible; consequently, in (Langton,84) an 8-state model for the self-reproduction of simple (yet non-trivial) non-universal automaton structures is presented in which the description of the automaton is stored in a dynamic "loop" (as opposed to static "tape"). Perrier et al. (1996) present a self-reproducing system supporting universal computation (but not universal construction) which is completely realizable and based on a synthesis of the approaches of von Neumann, Codd and Langton. However, they maintain that "the issue of trivial versus non-trivial self-reproduction is far from settled; as a simple counter-argument to Langton's viewpoint consider the observation that essentially any behaviour of a CA *ultimately* results from application of the cellular rule [emphasis added]." (p.337) Since the self-reproducing automaton is embedded in a CA, self-reproduction is *ultimately* reducible to the `physics' of the CA substrate.

It is worthwhile considering the nature of the relationship between computation- and construction-universality. Computationalists such as Langton (1989b) hold that there is a necessary link between the two as stated above; however, McMullin (1993a) argues that *computation* universality does not entail *construction* universality. As he states, "there are no such things as A_T-constructors or, more particularly, A_T-reproducers" (p.3), where A_T denotes a Turing machine (chapter 2). This does not mean that Turing-equivalent devices supporting universal construction are impossible; rather that Turing machines as originally conceived by Turing were not designed to support construction-universality. Correspondingly, McMullin holds that von Neumann's automaton was designed to demonstrate construction-universality; support for computation-universality was merely a *contingent* feature of the automaton and not essential to the proof. The conflation of computation- with construction-universality appears to have arisen from the fact that "von Neumann's attempted solution to [the problem of spontaneous, open-ended growth of complexity within automata] was heavily, and explicitly, influenced by

Turing's formulation of a certain formalised class of `computing machines'." (p.2) McMullin maintains that computation and construction constitute different functional (phenomenal, behavioural etc) kinds; moreover, functionalities which are substrate-relative: Universal-computation was defined in the context of Turing machines whereas universal-construction was defined in the context of CAs. However, if computation-universality can be defined independently of substrate, that is, if the Church-Turing thesis (chapter 2) is true, then a substrate supporting both universal-construction (such as CAs) and universal-computation becomes a logical possibility; a construction-universal substrate supporting the emergence of a computation-universal structure thereby becomes possible. McMullin alludes to this in stating that "it *can* ultimately prove useful to say something about the `computational' powers of *A*-constructors and/or their offspring" (p.5), whereby *A*-constructor is understood an automaton belonging to a class of automata, each of which is capable of automaton-construction. More than being merely useful, this is *necessary* if mind (consciousness, intentionality, intelligence etc) is emergent from life; the former can, for the sake of argument, be viewed as a computation-universal phenomenon arising in a construction-universal substrate which is itself computation-universal (assuming the Church-Turing thesis) although this latter property does not become manifest prior to the emergence of mind. Put simply, on this view, mind is *implicit* in life. (This important idea will be investigated further in chapters 6 and 7 in the context of a re-examination of the notion of emergence.)

Thus far, the discussion has focused on self-reproducing automata implemented in irreversible CA substrates, that is, in NRCAs. Toffoli (1977) has shown how non-trivial irreversible processes (such as universal computation and universal construction) can be implemented in a reversible substrate by constructively embedding a *n*-dimensional CA in a $n+1$-dimensional RCA ($n \geq 1$). (Morita et al. (1989) have proven the computation-universality of one-dimensional RCA.) Toffoli maintains that "if one concentrates one's attention only on [the embedded CA, a hyperplane in the embedding RCA substrate], one observes irreversible phenomena; the `information content', so to speak, of the process gradually decreases." (p.230) This focusing of attention *only* on the CA corresponds to the specification of an observation frame, and involves a measurement or `cut' of the RCA substrate (chapters 1 and 6); the irreversible CA (or NRCA) is thus, an abstraction or `virtual machine' embedded within the RCA substrate (section 5.5.9). While this approach allows for the embedding in real-time of a NRCA in a RCA, it is possible to simulate any NRCA using a RCA of the *same* dimensionality if the requirement for real-time simulation is relaxed (Morita,92). Importantly, as Morita et al. (1995) state, "from these results, existence of computation- and construction-universal (and thus self-reproducing) RCA can be concluded." (p.1) They present a scheme for implementing self-reproduction in a RCA using a variant of CAs called partitioned cellular automata (PCAs). Since PCAs are a subclass of CAs, it is conjectured that construction-universality in a RCA is possible. However, they correctly observe that "if we use these `emulation methods' to convert an irreversible CA to an RCA, a large amount of garbage signals are generated and spread out in reversible cellular space." (p.1) (As stated previously in section 5.5.6, these can usually be handled in RCAs by setting up some of the gates to act as `heat sinks' (Toffoli,80).)

5.5.7.2. Evolution

Arbib (1969) briefly discusses a conclusion which follows from a result in recursive function (or Turing-computability) theory due to Myhill, viz. finite programs contain the possibility of infinite improvements in successive generations of offspring without requiring any randomness in the mutations. Hence, a completely deterministic mechanism for biological evolution is possible and has been described in the context of the GOL, a non-reversible universal CA (or NRUCA) in section 5.5.3. Myhill's result is important since, as Laing (1988) has argued, it provides a means by which to *effectively* - that is, contingently or pragmatically, not logically or absolutely - overcome the limitations on automata which follow from Gödel's theorems (chapter 2). As Arbib (1988) states,

> while Gödel's incompleteness theorem points to an inevitable limitation of any axiomatization of arithmetic, Myhill's theorem points out the much less well-known fact that this limitation can be *effectively* overcome. And, of course, this process may be iterated mechanically again and again. (p.184)

Thus, by incorporating the possibility of deterministic variation and selection into the CA, it is logically possible to implement evolution in self-reproducing CA structures. As Berlekamp et al. (1982) state, an implication of evolution is the emergence of increasingly `intelligent' self-reproducing entities. Thus, CA-Life becomes the route to CA-Mind (assuming CTMi, that is, the computational theory of mind described in chapter 4).

5.5.7.3. Autonomy

Emmeche (1991) maintains that "in an autonomous living system, we cannot make the distinction between the entity being reproduced and an ultimate machine whose properties do not depend on the process of reproduction and which is not reproduced itself." (p.85) However, this view is problematic since a distinction *can* be made between the dynamic reproducing entity and the (assumed) static physico-chemical natural laws which provides the ontological substrate in which reproduction is realized. Hence, if a computational interpretation of physics (CA-matter) is possible (section 5.5.6), Emmeche's objection becomes untenable. A CA-like approach to modeling the emergence of autonomous self-reproducing structures has been explored by McMullin and Varela (1997) who report on a new computational implementation of autopoietic theory in the context of autocatalytic chemistry[42]. Other approaches include that of Pargelis (1996) and Koza (1993), the latter of whom describes a scheme for the

[42] This model is examined in further detail in chapter 6 in connection with a discussion of the link between autopoiesis, computationalism and mechanism.

"spontaneous emergence of self-replicating computer programs, sexual reproduction, and evolutionary self-improving behaviour among randomly created computer programs using a computationally incomplete set of logical functions." (p.259)

5.5.7.4. CA-Matter to CA-Life. How ?

If a computational or information-theoretic view of life (chapter 4) is adopted, then it follows necessarily that life is realizable in a UTM (or, equivalently, UCA). However, according to Fredkin (1990), life, in its most basic form, cannot be a RUCA phenomenon since it is clearly a non-reversible phenomenon (Prigogine,84). Fredkin asserts this despite Toffoli's (1977) proof that construction-universal (and hence, self-reproducing) structures can be embedded in RUCA. (Toffoli might argue that Fredkin and Prigogine's position rests on the adoption of a *local* perspective: On such a view total information regarding the automaton is impossible; hence, reversibility of life is also impossible[43]. By adopting a global or "God's Eye" view in which total systemic information *is* possible, the reversibility problem can be solved.) However, yet another problem concerns the emergence of CA-Life in a substrate which is a RUCA; since RCAs do not support self-organization (Wolfram,84b), how *can* CA-Life emerge from CA-Matter, and yet it *must* do if a unified framework of CEA is possible. Proposed solutions to this problem are examined in section 5.5.9.

5.5.8. CA-Mind

In the materialistic theory of mind, the sequence of thoughts in a rational process is causally determined not by the *logical* implications of the object about which we are thinking but simply by the *physical* changes occurring in the brain. This view immediately gives rise to the following problem: How can logic exist in or be supported by a universe which is ontologically material ? Logic exists at the conceptual level, yet materialism or *mechanism* (chapter 2) necessitates that secondary qualities (such as concepts) are grounded in primary qualities. However, explanations of the nature of this grounding are generally unsatisfying, thereby giving rise to the mind-body problem (chapter 4). Computationalists maintain that computationalism subsumes *both* the logical *and* physical aspects of naturalistic processes under the concept of computation; hence, the above problem does not arise[44]. Furthermore, in the context of the philosophy of

---

[43] In support of this contention, Margolus (1984) maintains that the embeddedness (or endophysicality) of observers within CA universes makes possible the perception of entropy increase: "from the point of view of creatures `living' inside an RCA, their inability to make use of complicated correlations between large numbers of cells means that for all practical purposes, the entropy of the automaton increases." (p.83)

[44] CA-computationalism supports both (1) the grounding of logic in physics, for example, Fredkin and Toffoli's billiard ball model of computation (Fredkin,82) and (2) the grounding of physics in logic, for example, the supervenience of digital mechanics on a RUCA substrate (Fredkin,90).

mind, Putnam (1988) has contested the viability of token-token[45] functionalism maintaining that

> physically possible sentient beings just come in too many `designs', physically and computationally speaking, for anything like `one computational state per propositional attitude' functionalism to be true. (p.84)

However, this position does not rule out type-type functionalism. As Putnam states,

> moving from the requirement that the `states' of speakers with the same reference (or believers with the same belief) be identical to the requirement that they be *equivalent under some equivalence relation which is itself computable, or at least definable in the language of computational theory plus physical science*, gives us enormous additional leeway [emphasis added]. (p.85)

Ths supports the notion of multiple-realizability (chapters 2 and 4) with respect to the correspondence (many-one) between brain states and mental states and is a strong indicator of surjectivity (Appendix). Surjectivity both implies and is implied by irreversibility; hence, in moving from CA-Matter to CA-Mind (via CA-Life), the RUCA substrate supporting CA-Matter must somehow give rise to a NRUCA. (This issue is examined in section 5.5.9.) Furthermore, surjectivity implies *self*-organization (chapter 3) and thus, the emergence of mind. As a corollary, and with respect to the issue of intentionality (chapter 4), it is possible that entities with the capacity for being informed, viz. CA-Minds, may *emerge* from CA-Matter; in computational terms, this can be interpreted as syntax (matter) giving rise to semantics (mind) with the latter as an *emergent property* (chapter 3) of the former. Additionally, interpreting intelligence and intentionality in behaviourist terms, that is, by adopting the *intentional stance* (Dennett,85), CA-Mind can be realized independently of the postulated emergence of consciousness.

There are no theoretical reasons for the impossibility of CA-Mind given the assumption that intelligence (as the determining characteristic of mind) is a computational phenomenon; indeed, this position is consistent with the classical view of AI (chapter 4). Furthermore, CAs may be shown to provide support for the possibility of "strong" AI (chapter 4) once it is recognised that they are formally (computationally) equivalent to neural networks and semantic networks (Farmer,90). Although it is uncommon to find examples of CA-based models of the mind or brain in the literature, there have been notable counter-examples: For example, Wuensche (1993) has modelled memory and cognition using CAs and random boolean networks and computational modeling of microtubular function within cytoskeletal lattices using CAs was first reported in (Smith,84). On the basis of this latter idea, Hameroff et al. (1988) have developed a series of CA-based models of microtubular action in the brain, viz. microtubular automata or MTAs. As they state,

---

45 That is, an isomorphism, or one-one correspondence, between brain states and mental states.

> biological activities including guidance and movement of single-celled organisms, transport of molecules within cells, and cognitive functions of the human brain may utilize MTA behaviour .. In nerve cell axons and dendrites, massively parallel arrays of interconnected MT and neurofilaments could serve as specialized information circuits. MT automata gliders (8 to 800 meters/sec) may exist as traveling depolarization waves, solitons, or localized electrostatic fields which bind and transport molecules. In nerve cells, traveling gliders may correlate with traveling nerve membrane action potentials. Nerve membrane depolarization may interact with MT automata by transiently altering dipole coupling thresholds via ionic fluxes, voltage gradients or direct connection to the cytoskeleton. This coupling of the brain/mind as a hierarchy of nested automata with a previously overlooked basal dimension. (p.543)

Appealing to quantum non-locality (chapter 4) in order to explain the `unitary sense of self', non-deterministic free-will and non-algorithmic `intuitive' processing, Hameroff (1994) maintains that cytoskeletal microtubules within neurons could be a possible site for quantum effects. Specifically, "consciousness may emerge as a macroscopic quantum state from a critical level of coherence of quantum-level events in and around a specific class of neurobiological microstructure: cytoskeletal microtubules within neurons throughout the brain." (p.92) This leads Hameroff to assert that

> brain-wide microtubule-based quantum states, when coupled to synaptic events and neuronal firing, can help account for [various] properties of consciousness (pp.93-94)

Thus, consciousness in CA-Minds, according to Hameroff, is taken to be an emergent phenomenon. This approach is consistent with a recognition of the "hard problem" of consciousness (that is, experience, ontological subjectivity, first-personhood or what-it-is-likeness) described in chapters 4 and 7[46], a problem which is not adequately resolved on the behaviourist interpretation of mind.

5.5.9. Emergence and Embedding

Svozil (1993) provides a clear statement of the role of automata such as CAs under the finite nature thesis in the production of artificial realities, viz.

> every automaton is a universe of its own, with a specific `flavour', if you like. Programmers may create `*cyberspaces*' (a synonym for automaton universes) of their imagination which are, to a certain extent, not limited by the exterior physical laws to which they and their hardware obey. Seen as isolated universes, some of these animations might have nothing to do with our physical world. Yet, others may serve as excellent playgrounds for the physicist. (p.22)

Additionally, Davies (1992) maintains that if physics is identical to computation, then the universe becomes not only equivalent to, but *identical* with its own simulation, that is, the universe becomes its own emulation (chapter 4). These two statements are

[46] As stated previously in chapter 4, the "hard" problem (Chalmers,96) is slightly different to the category problem (chapter 7) in that the latter demands an *emergentist* solution to the problem of the relation between ontological objectivity and subjectivity.

important since they assert that the universe can both simulate (Svozil) and be simulated (Davies) by itself, entailing the embedding of universes and the construction (via design or emergence) of virtual machine hierarchies (chapter 4). The notion of CA embedding has already been discussed in connection with the embedding of construction- and computation-universal NRUCAs in RUCAs in section 5.5.7. Additionally, it is possible to embed RUCAs in NRUCA substrates; this is readily demonstrated by the fact that RUCA universes can be simulated using nonreversible programs (software) executing on nonreversible physical computers (hardware). NRUCAs can simulate[47] RUCAs and visa versa since both are instances of UCAs which are, in turn, CA realizations of UTMs. This equivalence is graphically represented in Fig 5.8:

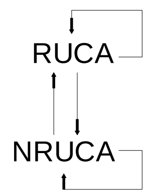


Fig 5.8 Equivalence of RUCAs and NRUCAs.

Moreover, since software and hardware are mutually interchangeable - a basic result arising from the duality of computing systems (chapters 2 and 4) - virtual worlds (machines) can be nested (embedded) to arbitrary depth as shown in Fig 5.9:

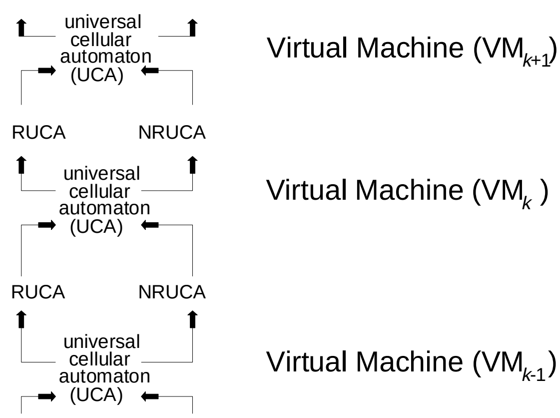


Fig 5.9 Nesting of virtual machines. (Branching indicates a choice of UCA.)

According to Tipler (1996),

> Turing machines can emulate other Turing machines. In fact, there is one Turing machine, called the universal Turing machine, which can emulate all Turing machines, including itself. We can thus have a hierarchy of machines emulating other machines. Turing machine T_0 may be a real machine, but

47 In fact, *emulate* since both simulat*or* and simulat*ee* are of the same ontical kind, viz. computational universes.

> inside it there is a virtual Turing machine T_1, and inside T_1 there is a virtual machine T_2, which in turn codes virtual machine T_3, and so on. These levels of virtual machines inside virtual machines are called *levels of implementation*. The higher-level machines operate completely obliviously of the machines on lower levels. For nothing happens to the higher levels if one or more of the lower level machine are replaced by other completely different machine - provided, needless to say, that the replacement machines are capable of emulating the higher-level machines at the same speed. As a general rule in computers, machines from various levels do not mix, but this is done to simplify life for human computer programmers rather than something required by the mathematics. If a machine is transferred to a higher level of implementation, it is said to be *uploaded*, and if it is transferred to a lower level of implementation, it is said to be *downloaded*. (pp.36-37)

It should be noted from the above statement that a primitive or ontical (substrate) level of Turing machine or CA is necessary in order to ground the hierarchy; thus, the embedding is unidirectional rather than bidirectional (as was implied in fig 5.9). (Ali (1998a) describes a bidirectionally infinite scheme for embedding CAs in which the grounding problem is solved by postulating observers capable of `cutting' the hierarchy, thereby establishing a substrate and laws of physics governing its behaviour relative to the observers.) The role of T_0 (or, equivalently, CA_0) is crucial: it is responsible for dynamically realizing or actualizing the static embedded virtual machine potentialities; in short, T_0 (or CA_0) is a necessary condition for executing the essentially Platonic entities embedded within it, converting them from passive programs into active processes (chapter 2). This is an extremely important point which will be addressed further in chapter 6.

5.5.9.1. Conditions For Embedding

Toffoli (1978) describes some of the requirements for embedding systemic phenomena within computational media such as CAs:

> In general, the embedding of a particular guest system, or *object*, in an assigned host system, or *medium*, is achieved by suitably constraining the initial conditions of the medium. In this way, *both parameters and state variables* of the object system are mapped into *state variables* of the medium. (p.395)

This procedure is equivalent to *microprogramming*. Furthermore, assuming a uniform or isotropic medium,

> the embedding should be realized in such a way as to prevent undue interference on the objects behavior on the part of the surrounding portion of the medium or *environment*. (p.396)

Toffoli (1978) makes an important point regarding the possibility of support for non-reversible phenomena in reversible environments, viz. "an embedded object or observer can perform irreversible tasks only at the cost of gradually contaminating the surrounding environment with computation by-products - the second law of thermodynamics still holds - while in an irreversible medium the necessary negative entropy can be extracted, so to speak, from the irreversibility of the medium itself. (As a consequence, the modelling of irreversible systems in a reversible medium requires that the medium be infinitely extended)." (p.399) Consequently, Toffoli (1977) maintains that a necessary condition for embedding stable NRCAs is that the RCA substrate must be propagation-unbounded[48]. (However, this conflicts with the second axiom in Fredkin's (1990) finite nature thesis (chaper 2), viz. finitude of the universe.) An additional requirement is identified by Toffoli as follows:

> In order to identify a particular embedded system, or *object*, in a given *medium*, one has to specify a `portion' of the medium itself, associated with the object's *extension*, and initial conditions or other constraints for this portion, associated with the object's *structure*. (p.397)

Specifying a `portion' of the medium necessitates the drawing of a distinction (Spencer-Brown,69) - a `cutting' (chapter 6) - *within* the medium; alternatively, it can be viewed as necessitating *measurement* (Pattee,95), thereby involving the act of observation (Cariani,89). Hence, if a CA universe is to be self-contained, it must support the emergence of *intrinsic* observers (chapter 3). According to Toffoli (1977), "a model capable of supporting a wide range of computing and constructing capabilities makes it possible to explicitly characterize the observer and its interaction with the observed system." (p.214) Hence, for Toffoli, observation can be defined in objective terms[49]. This position is supported by Adami et al. (1996) who define information as the mutual entropy between two systems while maintaining that

> information is *not* a list of symbols [for] without referring to an environment that a string is to be interpreted within, the notion of information is meaningless. (p.7)

They further maintain that in the context of reversible Turing machines (chapter 2) with finite tapes (representing physical universes), any strings computed from the tapes correspond to measurements performed on the abstract universe; hence, measurement *is* reducible to computation. The reduction is supported by the unification of thermodynamics, information theory and computation theory under the concept of entropy (chapter 4). What is problematic with this position is that no explanation as to *why* a measurement occurs is forthcoming. Computation, according to Adami, is equivalent to symbol-manipulation; furthermore, any physical manipulation of physical objects can be *interpreted* computationally once such objects are assigned symbolic status. However, Cariani (1989) contests the postulating of computational processes, that is, symbolic manipulation of primitive observables, in the absence of an *observational*

---

48 Given the global CA map τ, the *propagation* of a configuration c is the sequence $\langle \tau^t c | t=0..\infty \rangle$. A propagation is *bounded* if the corresponding sequence of configuration diameters is bounded. A CA is *propagation-unbounded* if every configuration except the blank one has an unbounded configuration.

49 Searle (1992) has implicitly contested this position in asserting that syntax is not intrinsic to physics. On his view, systemic observables cannot be defined independently of the intentionality of the observer and intentionality is grounded in consciousness. Hence, an objectivist account of observation is considered problematic.

reference frame. Thus, it is meaningless to speak of observation or measurement in the absence of observers. This line of critique can be extended further: is it meaningful to speak of observations in the absence of *conscious* observers ? The point is well made by Eddington (1928) who maintains that

> although a change described as sorting is the exact opposite to a change described as shuffling we cannot imagine a cause of sorting to be the exact opposite of a cause of shuffling. Thus a reversal of the time-direction which turns shuffling into sorting does not make the appropriate transformations of their causes. Shuffling can have inorganic causes, but *sorting is the prerogative of mind or instinct*. We cannot believe that it is merely an orientation with respect to the time-direction which differentiates us from inorganic nature [emphasis added]. (p.99)

5.5.9.2. The RCA→NRCA Problem

Computational analogues of the phenomena associated with classical and quantum physics are realizable in RCAs (section 5.5.6); however, life is an irreversible phenomenon. Although, CA-Life (a NRCA) can be embedded in CA-Matter (a RUCA), if CEA is true then this embedding must arise spontaneously; that is, a RCA→NRCA transition (Ali,94b) must occur. However, Bennett and Landauer (1985) maintain that

> since the output is implicit in the input, no [reversible] computation ever generates information. (p.38)

Thus, computational or deductive emergence (chapter 3) - which, *by definition*, involves the production of information (Baas,93) - is not supported by reversible computational systems such as RUCAs[50]. This result is consistent with Polanyi's (1966) interpretation of emergence in which it is asserted that a system cannot define its own boundary conditions (chapter 3). However, specification of such conditions is necessary for demarcating structures (such as living organizations). In RCAs they are of *necessity* non-emergent and must be designed *into* the system, externally preset as initial conditions. Wolfram (1984b) maintains that "self-organizing behaviour occurs by virtue of the irreversibility of cellular automaton evolution" (p.34) and that "evolution to attractors from arbitrary initial states allows for `self-organizing' behaviour in which structure may evolve at large times from structureless initial states. The nature of the attractors determines the form and extent of such structures." (p.3) Thus, as stated previously, irreversibility is a *necessary* condition for self-organization and, by association, emergence. The distinction between NRCAs and RCAs with respect to their respective capacities for emergence is shown in Table 5.9 (overleaf).

---

50 It is assumed for the sake of present argument that irreversible computation is not an *a priori* infomationally-closed phenomenon; this is true in NRCAs when the size of the lattice approaches infinity and the CA rule is complex or class IV (chapter 2).

NRCAs	*RCAs*
surjective self-organizing (or emergent) `open' systems (negentropic)	injective non-self-organizing (or non-emergent) `closed' systems (entropic)

Table 5.9 Comparison of NRCA and RCA properties.

Surjectivity contributes to the stability (robustness) of NRCA structures; for example, structural perturbations (or "noise") caused by collisions with other CA structures are tolerated to a much greater extent than is the case in functionally isomorphic RCAs. This is important because life is a relatively robust phenomenon (chapter 4). However, NRCAs are still extremely brittle. As Wolfram (1984b) states,

> the fraction of configurations which may be reached after one time step in cellular automaton evolution, and which are therefore not on the periphery of the state transition diagram, gives a simple measure of irreversibility. (p.4)

This point can be made more formally as follows. Let s_d denote the total number of nodes (states) at a depth d in the state-transition graph for an attractor basin in the basin of attraction field associated with a NRCA. (Depth is measured from the attractor to nodes on its basin; d=0 denotes the attractor.) Let s_k denote the number of nodes converging on a parent node at depth d-1, where $d \geq 1$. s_k/s_d can be used as a *naive*[51] measure of the robustness of the parent node at d-1. (Attractors are maximally robust, that is, s_k/s_d=1, with respect to each basin of attraction in the basin of attraction field.) As stated above, RCAs (that is, injective automata) are even less robust than NRCAs. As Toffoli (1980) states,

> computation [and by association, construction] in reversible systems requires a higher degree of "predictability" about the environment's initial conditions than computation in nonreversible ones. (p.643)

For this reason, Landauer (1967) argues against reversible computation:

> Information processing inevitably has to lose information .. If we do not lose information, we are handling numbers, in a calculation which has been rigidly forseen, step by step. In a general-purpose computational process, we must be able to do more than that; we must have the ability to make decisions during the computation. This is turn consists of the ability to take intermediate computational results and from them, rather than from all the intricate history leading up to these signals, proceed further. *If we do not throw away information about the history, a given logical signal will depend on the exact way it was reached; then, after a sufficient number of steps, the accumulation of these historical characteristics would lead to errors* .. [A] standardization process [is used to eliminate] the earlier unnecessary history [emphasis added]. (p.107)

51 `Naive' since the structure (topology, geometry etc) of the CA state-space has not been taken into consideration in the definition of the metric.

`Standardization' implies measurement, a necessarily statistical process involving the selection and quantification of certain characteristics to the exclusion of others (Pattee,95). (Clearly, this is a form of `cutting' as will be argued in chapter 6.) This leads Wolfram (1985) to propose the following solution to the RCA→NRCA problem. He identifies two necessary conditions for applying thermodynamic principles to CAs, viz. (1) reversibility (invertibility) and (2) "coarse-graining". As he states,

> to apply thermodynamics one must also `coarse-grain' the system, grouping together many microscopically-different states to mimic the effect of imprecise measurement. Coarse-graining in cellular automata may be achieved by applying an irreversible transformation, perhaps a cellular automaton rule, to the cellular automata configurations. A simple example would be to map the value of every other site to zero.

Again, the problem with this scheme is that a graining-operator cannot spontaneously emerge from the CA substrate since it is a RCA and hence, does not support self-organization. For this reason, Cariani (1989) maintains that measurement (a non-computational process on his scheme) constitutes an irreducible phenomenon involving the specification of an observational reference frame by an endophysical (intra-systemic) observer, that is, an observer situated *within* the universe. On this basis, computationalists such as Crutchfield (1994) have argued analogously in favour of intrinsic (that is, endophysical) emergence with CA-type systems. However, the measurement problem arises since computational emergence necessitates specification of *Obs* (Baas,93), that is, an observational reference frame (chapters 3 and 6) which, on CEA, must itself be emergent.

An alternative approach which retains substrate reversibility (and hence, is consistent with CA-Matter) while simultaneously supporting self-organization is implicit in the following statement. Svozil (1990) maintains that "the capability of computable functions to `produce' uncomputable output on computable initial values may have some far-reaching consequences in the physical perception of reversibility .. [For example,] the creation of algorithmic complexity may be perceived as a formal aspect of irreversibility." (p.425) On this basis, he goes on to offer the following conjecture:

> a .. technically rather subtle and speculative contribution to an arrow of time would be a reversible CA whose time evolution is recursively enumerable [or computable] but whose inverse is uncomputable. This would correspond to a computable bijection $f(x)=y$ whose inverse $f^{-1}(y)$ is uncomputable. (p.426)

According to Toffoli and Margolus (1990), the dynamics of a CA will be invertible (reversible) if the state-transition function is invertible. However, can an invertible CA (RCA) be embedded in a non-invertible substrate (NRUCA) such that the dynamics of the invertible CA `virtual machine' in the backward-time direction are non-invertible ? This question can be answered in the affirmative *if* non-invertibility or backwards non-computability is a consequence of *lattice* directionality as contrasted with state-transition rule directionality (or surjectivity). Thus, Petrov (1996,1997) presents an extension of the basic CA formalism, viz. *CAs on infinite homogeneous directed graphs*, in which it is possible to obtain asymmetrical configurations from a single non-blank cell in the initial configuration, using perfectly symmetrical rules and homogeneous space. The significance of this scheme is that it provides a means by which to explain the origin of asymmetry in a homogeneous universe with symmetrical laws. In support of this framework in a *physical* context, Petrov (1996) states that

> there are no serious reasons to state that homogeneous space has to be modeled only by lattice structures; according to us every topologically homogeneous cellular structure could be a worthy candidate for a discrete mathematical model of any physical phenomenon. (p.279)

Additionally, in support of this framework in a *biological* context, Petrov (1996) argues against the equating of non-trivial reproduction with self-reproduction, viz.

> although the process of reproduction of living organisms always leads to the appearance of same-of-kind, it is more a process of *resemblance* than an act of exact duplication of the prime source. The preservation of some parental attributes and the mutation of others is the basis of the improvement of species and of the appearance of new ones [emphasis added]. (p.280)

Petrov identifies the "Big Bang" with a Garden of Eden (GOE) state and explains the "arrow of time" as a consequence of the injectivity (reversibility) of the underlying finite CA. (Importantly, the CA that is the universe must be injective or reversible and not bijective or invertible in order that GOE configurations are possible.) However, the asymmetry is non-emergent in the "strong" sense implied by *creatio ex nihilo* (chapter 6), that is, a genuinely new feature or characteristic of the universe since although the vertices in the digraph are indiscernible, the edges are discernible and directionally-labelled and it is this orientation which accounts for the asymmetry. Thus, irreversibility is, in fact, artifactual, that is, non-emergent in the strong sense since the directionality of evolution is a manifestation of the directionality of the lattice. However, it should be recognized at the outset that Petrov uses the term invertibility in a slightly different sense to Fredkin. As he states (1997),

> the problem is that what we call `reversible CAs' are, in fact, `properly injective CAs on finite configurations only'. These should be defined as CAs for which every finite configuration has not more than one (finite) predecessor, and it is important to add that there exist some configurations, which under these rules have no predecessor at all (`Garden-of-Eden' configurations). To make things clear, what we call `invertibility', is equivalent to `bijectivity over the set of all finite configurations only'.

Furthermore, thus far it has only been shown how trivial reversible rules (such as XOR) can give rise to irreversible structures; it has not been shown how non-trivial reversible rules can generate universal irreversible structures (such as computers and constructors). However, this approach offers one of the most promising routes to the realization of CEA under a CA substrate.

5.6. CA-Computationalism: Beyond Alexander

CAs provide the means by which to solve a number of outstanding problems associated with Alexander's metaphysics.

5.6.1. Logical Necessity

Collingwood (1945) has criticised Alexander's ontology on the grounds that its strict immanentist approach to categories - a consequence of adopting a naive empiricism - fails to establish grounds for necessary causal connection between spatio-temporal events:

> It looks as if Alexander, deeply influenced by Kant but resolving at all costs to avoid Kant's subjectivism, had cut out the Kantian [transcendental] categories altogether, because they are merely subjective necessities of thought, and contented himself with the [immanent and empirically-grounded] schemata by themselves. But if you cut out the category of cause and substitute its schema, you are cutting out the idea of necessary connection and trying to content yourself with mere uniform succession. (p.163)

Furthermore,

> if the method of philosophy is purely empirical, if the universal means merely the pervasive, the necessary merely the actual, thought merely observation, a system built on this method can have in it no driving force or continuity; there is an element of arbitrariness in every transition, and a reader who stubbornly asks, `*why* should space-time generate matter; *why* should matter generate life; *why* should life generate mind ?' and so on, will get no answer; he will only be told that he must not ask such questions but must accept the facts in a spirit of natural piety. Yet if the child is the father of the man, surely the first duty of natural piety is to respect, and endeavour to satisfy, the childish tendency to ask questions beginning with *why*. (p.163)

CA-computationalism solves the above problem by implementing a variant of formism (chapter 2): At the substrate level, the *local* (FSM) state-transition function in a CA is fixed (static) and deterministic; derivatively, the *global* (CA) state-transition function or basin of attraction field is static and implicit in the functional connectivity of the FSMs, the FSM state-transition function and the initial state configuration of the CA. Importantly, the formism of CAs is not in the patterns, but in the `laws of form' which are responsible for pattern generation. This position is consistent with Alexander's internalist view of relations. As Brettschneider (1964) states, "on Alexander's grounds, relations cannot have the status of logical subsistents [that is, Platonic forms] because they go to make up an individual. The structure of the individual does not permit connection by abstract or formal relations." (p.46)

5.6.2. Explanation of Emergence

Brettschneider (1964) maintains that "in his discussion of the transitional stage between infinite and finite motion [the former being associated with Space-Time and the latter with complexes of space-time], Alexander seems to advocate an hypothesis of spontaneous generation. It is quite in order to infer that he is talking about *the genesis of something from nothing* [emphasis added]." (p.60) However, if an idealist interpretation of Alexander is adopted, in which components within the metaphysical system are internally as opposed to externally related (that is, if an ontological coherence-theoretical as opposed to correspondence-theoretical position is adopted), the creation problem is resolved since "the difference between infinite and finite is of degree, not kind." (pp.60-61) Furthermore,

> organization as a categorial property is immanent in pure Space-Time. Organization is possible because the inherent restlessness of Time causes Space-Time to restructure itself constantly .. The elements and conditions of cosmogenesis are present in pure Space-Time. There are no additives. (p.62)

However, this solution is purely *descriptive*. As Brettschneider states, "on these grounds philosophy may only describe the process of creative evolution as a *fait accompli*. We may neither proffer explanations nor legislate for the universe .. Emergence is a fact of nature for Alexander; its existence [or rather, its occurrence] cannot be denied; its awesomeness induces the attitude of `natural piety'." (p.61) The important and controversial issue of *creatio ex nihilo* (creation from nothing) which is intimately bound up with discussions relating to the maxim *ex nihilo nihil fit* (from nothing comes nothing) is examined in chapter 6 in connection with a preliminary investigation of the relations between Being and beings under *poiēsis* (coming-forth or becoming). Again, CA-computationalism solves this problem through a variant of Platonism at the level of CA-Matter, viz. by postulating static and deterministic state-transition functions as the "laws of physics" in CAs. Thus, CA-computationalism is consistent with an interpretation of emergence as the `unfolding' of form *implicit* in the structure of the CA. (A similar position advanced by Simon (1981) was described in chapters 2 and 4.)

5.6.3. Identity as Empirical

Alexander (1920) states that

> what changes are compatible with the retention of substantial identity is an *empirical* question which can only be decided by reference to each case or kind of cases. (Vol.I, p.273)

As will be argued in chapter 6, where the notion of `cutting' first introduced in chapter 1 is examined in further detail, categorial identity depends on the way in which categories are established, that is, where categorial boundaries are drawn. Again, the issue of forms (whether viewed as Platonic or as conceptual, and thereby, anthropocentric) arises and again, CA-computationalism provides a solution to this problem in terms of the unfolding of stable structures in the basin of attraction field.

5.6.4. A CA model of Alexanderian Space-Time Incorporating Reversibility

The basis for constructing a CA model of Alexanderian Space-Time is contained in the work of Ilachinski (1987) and Halpern (1989, 1990). As Halpern (1989) states, in standard CAs "only the values of the sites are altered. The lattice structure itself is unaffected by these rules." (p.405) Halpern presents an extension of the standard CA formalism in which rules defining the evolution of the lattice structure are incorporated; hence, *Topological Automata* (TAs), in which *both* site values *and* site connectivities (or links) evolve. This extended CA formalism can itself be modified such that *only* the link-patterns evolve as a function of time; thus, CA space-time evolution can be defined purely in terms of changes in the pattern of connections between cells. Redefining the `state' of a cell in terms of the *degree* of its connectivity to other cells provides support for a non-Newtonian (or non-`container') interpretation of CAs and is consistent with Alexander's (and Wheeler's) conception of geometrical spacetime as the primal-stuff of reality. Defining cell-state in terms of connectivity allows for the reduction of the (cell-state, cell-link) dualism to a geometrical monism and thereby eliminates a degree of freedom (arbitrariness) associated with TAs. Furthermore, Fredkin's second order technique (section 5.5.5) can be used to generate a RCA version of the Alexanderian TA and it is conceivable that Petrov's (1996) digraph approach can be incorporated so as to support self-organization in a reversible TA substrate.