Relativistic Ontologies, Self-Organization, Autopoiesis, and Artificial Life: A Progression in the Science of the Autonomous

Part II — A Scientific Development

David Vernon* and Dermot Furlong
School of Engineering
Trinity College Dublin
Ireland

Abstract

In a sister paper, we have looked at the philosophical aspects of the development of autonomous systems, touching upon the inadequacies of conventional (positivistic) ontologies and philosophies of science, and we have described an alternative relativistic ontology. We argued that self-organization is a necessary condition for autonomous systems and we highlighted the difficulties that this raises for conventional representational approaches to autonomous systems. We discussed a methodology for discourse in relativistic ontology (*Systematics*) and, based on this, we argued in favour of a spectrum of autonomy. In this paper, we try to show how autopoiesis can be interpreted as a particular instance of autonomy in this spectrum. We now proceed to describe the progress which has been made towards the development of a computational simulation of autopoietic organization, beginning with a formulation in terms of the Calculus of Indications (incorporating Varela's extensions to include autonomous forms), and incorporating the *Systematic* formulation.

1 Designing Autonomous Systems

In a companion paper, we developed a well-founded characterization of autonomous systems, from a philosophical standpoint. This characterization is grounded upon the acceptance of a relativistic ontology — positing a spectrum of existence and being — wherein autonomy is seen as a characteristic

*David Vernon is presently on leave at the Commission of the European Community.

of a system above a specific level of being. This level is 'marked' by the organization exhibited by the entity and autonomy arises once a critical level of self-organization is achieved. Our task now is to take this philosophical characterization and develop it into a scientifically-useful theory. To do this, we will first make two diversions from our route to this goal, diversions which will ultimately return us to our path better equipped to achieve our objective.

1.1 Autopoiesis

In 1979, two scientists — Francisco Varela and Humberto Maturana — postulated that all living systems share a common organisation which we implicitly recognise by calling them living[13, 9]. In Varela's words: 'a living system is defined by its organisation [and] ... it can be explained ... in terms of relations, not component properties'. Thus, Varela and Maturana too advocate a shift in emphasis from the study of the structure of instantiated systems to the organisation underlying and facilitating the structure in the first place. Instead of investigating the behaviour of systems exhibiting autonomy and the concrete biological and computational 'implementation' of this autonomy, the study addresses the reason why such behaviour is exhibited in the first place. That is, Varela and Maturana advocate the of study systems and, specifically, of the abstract organisation of systems.

If one accepts the reasonable axiom that living systems exhibit fundamental characteristics of autonomy and unity, then what is needed is an explicit and unambiguous definition of the type of organisation which facilitates that unitary autonomy. According to Varela and Maturana, a living machine's stability is accounted for by processes which are internal to the system. Thus all feedback is internal to the machine or system and all references for this feed-back is internal to the system. This arises through the mutual interconnection of processes (i.e. its organisation), and the mutual interdependence of processes, again its organisation. This organization, then, is the source of the unity and autonomy of the system and exactly the requirements we make of, and observe in, the characteristics of living systems. Thus, such systems actually distinguish themselves (set themselves apart) from their environment through this organisational self-specification and self-production. Maturana coined word *autopoiesis* to convey this organisation (literally from the Greek for self-producing) and it is worth quoting here Varela's definition of autopoiesis[13]:

> *An autopoietic system is organised (defined as a unity) as a network of processes of production (transformation and destruction) of components that produces components that: (1) through their interactions and transformations continuously regenerate and realise the network of processes (relations) that produce them; and (2) constitute it (the machine) as a concrete unity in space in which they exist by specifying the topological domain of its realization as such a net-*

work.

We see that that it is the organisation of the system which is of primary importance and that this organisation refers to the relations of components, not to the components themselves. These components through their interactions continuously regenerate the relations, i.e., the network of processes. An autopoietic system is, thus, a self-referential system. The importance of self-reference is paramount for two reasons: it is the means by which the autopoietic system manages to 'specify itself', thus identifying its unity and autonomy. Secondly, it is the main reason why a reductionist analysis of living systems (which is the result of the pervasive positivistic world view we talked about above) fails to capture the complete essence of the system since reductionism cannot, by definition, accomodate the circular nature of such a self-referential definition.

Varela himself lists four consequences of autopoietic organisation and it is of value to reproduce them here.

1. *Autonomy.* Autopoietic machines are autonomous, i.e., they subordinate all changes to the maintenance of their own organisation (and identity). Thus, self-preservation is the most fundamental characteristic of autopoietic systems.

2. *Identity.* Autopoietic machines have individuality in that, in keeping their organisational invariance (though in constant re-production), an active identity is maintained despite interaction with that which is not the machine, e.g., an observer or the environment.

3. Unity. Autopoietic machines are unities, since its boundaries (i.e. itself) are self-defined, by virtue of the system's specific organisation.

4. *Closure.* Autopoietic machines have neither inputs nor outputs. And, as a consequence of the last point above, they are thus organisationally-closed systems. However, they can be perturbed by independent events (in the environment) and such perturbations may cause further perturbations of the environment, by the system.

Clearly, this is a constructivist philosphy which is very close to, if not identical with, that which we discussed in the previous chapter. Maturana's and Varela's working hypothesis is that living systems (and, by extension, artificial simulations of them) are autonomous stable homeostatic systemic complexes (unitary entities) whose primary function is to maintain their stable autonomy within a domain (a universe) of continually changing perturbations. The necessarily dynamic nature of such systems requires interaction with, and intrusion into, the environment (that which is not the system) with attendent change in the nature of the environmental perturbation of the system. The systemic complex is a unitary system, which is part of the universe, but which sets itself apart and becomes distinct from the universe (or environment) through its autonomic organisational behaviour, and thereafter is concerned wholly with the dynamic maintenance of this 'distinction', while continuously being perturbed by the environment and while, in turn, perturbing the environment.

This definition of autopoietic systems allows an interesting interpretation, or reassessment, of the concepts of cognition and perception. If we accept that living systems are, in fact, autopoietic machines then consequently cognition may be viewed as the activity of any autopoietic system which is concerned with the maintenance of its autonomy, i.e., its autopoiesis. This broad definition of cognition very useful, for if we are able to model the machine using autopoiesis, identifying the network of productions of components that characterize the system, then we have a very well-defined statement of what it is for that machine to engage in cognitive activity. Cognition is, in fact, the manner in which an autonomous unit maintains its autonomy in the face of continuous perturbations from its environment and perception, an aspect of cognition, is that systemic activity which facilitates the preservation of its autopoiesis. Since the overriding priority of the autopoietic system is to maintain its autonomy, cognition is thus an activity which facilitates the autonomy; making coherent *sense* of the flux of environmental perturbations, such that this sense, this interpretation, further facititates the systems autonomy.

Stated boldly, *cognition is a concept which is instantiated to explain (stand for) an underlying dynamic interplay of components of the system, the flux of interrelations, evident in living (autopoietic) systems. Perception is identically the systemic activity which is manifested by continuous interaction (intrusion/being intruded upon) with that which is not the unitary system.*

1.2 The Calculus of Indications

At this point, we have a novel, interesting, and useful characterisation of autonomous systems, in general, and of cognitive and perceptual systems, in particular. We have developed a philosophical and ontological understanding of autonomous systems, we have a 'Systematics' which enables us to identify the boundary conditions of autonomy, and now we have a definition of the organizational requirements of autonomous systems. However, we have no tools as yet for formally investigating and manipulating autopoietic mechanics, if we may put it that way. What is required is a formalism or calculus which will allow us to discuss qualitative phenomena (for of such is our subjective world of thought comprised) in a quantitative manner without recourse to numeric calculi (for an elegant demonstration of the inadequacy of numeric calculi in this domain of discourse, see [4]).

In the late 1960's when G. Spencer-Brown introduced the Calculus of Indications in his book *The Laws of Form*[12] and Varela has adopted, and adapted, this calculus to express the dynamics of autopoiesis. Spencer-Brown's calculus is non-numerical and it is based on the most elementary of all conceivable operations: that of ‘making a distinction’ — of indicating something. For example, one might ‘indicate’ a chair, in which case one is presented with a universe or domain in which a chair is distinguished; there is the chair and implicitly there is everything else. This is an intuitively pleasing approach as it is an act that we as living systems are continuously performing. As such, the calculus is single-valued: indicating ‘not a chair’ is not the inverse of making the distinction of a chair in the first place. It is an entirely different distinction. (In fact, the ‘inverse’ is to indicate the indication of the chair, which is, in effect, to make no distinction at all). Thus, it can be seen how this monadic system obviously differs from the much more familiar dyadic system of conventional ‘two-valued’ logic. Nonetheless, conventional Boolean logic is encompassed in the Calculus of Indications and its algebra is sufficient to perform all of the operations of Boolean algebra.

One of Spencer-Brown’s central theses is that distinction (which is an active concept), giving rise to an indication, is a fundamental notion and that all indications are, at this level, identical. Thus, one can deal with fundamental notions of a primary distinction and an attendant indicational space. The act of distinction is abstracted from particular domains and the embodiment of pure general indicational space (and Calculus of Indications) is posited.

Given a space (e.g. the world) and give a distinction in that space (e.g. a chair), the parts shaped by the distinction are the states of the distinction. The space and the states, *together*, are the *form of the distinction.*

States which are distinguished by the distinction are signified by a mark of distinction $\rceil$. This, then, is the marked state. $\rceil$ is also referred to as a *cross.* The former interpretation is static and allows the construction of forms; the latter interpretation is dynamic and allows the embodiments of activity in distinction. The state not marked with a mark is called the unmarked state; thus, there is only one explicit symbol in this arithmetic, that is $\rceil$ the cross or mark.

Any arrangement of marks (and non-marks) considered together with regard to one another, i.e. as part of the one form, is an expression. The state indicated by an expression is the value of the expression. $\rceil$ and   (i.e. no mark) are expressions. They are referred to as simple expressions. There are no other simple expressions.

Recalling the dual interpretations of $\rceil$ as either static form or dynamic distiction, we see that any expression is dually a (perhaps complex) distinction or value and a dynamic inter-relationship of mutual indication. This duality of the static and the dynamic is central to the usefulness and beauty of the calculus of indications.

There are two axioms in the calculus.

Axiom 1: The form of condensation.

$$\rceil\rceil = \rceil$$

This is interpreted as follows. If two crosses are contained in the same space, their value is that of distinguishing twice and, thus, their value is that of the marked state.

Axiom 2: The form of cancellation. $\overline{\overline{\;\;}|}| =$

This is interpreted as follows: if a cross crosses another, we undo the distinction; hence the value is that of the unmarked state.

Note that the mark implicitly embodies the idea of a distinction between (a duality between) the inside and the outside; that which is 'covered' by the cross being the inside. This is particularly elegant when one proceeds to consider complex indicational expressions (forms) such as:

$$\overline{\overline{\overline{\;\;}|\,\overline{\;\;}|}|\,\overline{\;\;}|}|$$

The overhang of the cross identifies the extent of the marked state.

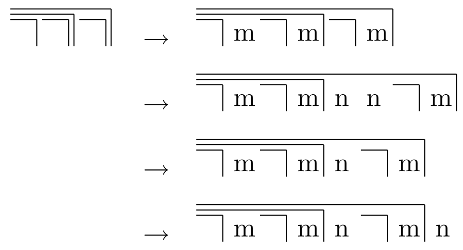
An expression such as this can be simplified by repeated application of the two axioms. Thus:

$\overline{\overline{\overline{\;\;}|\,\overline{\;\;}|}|\,\overline{\;\;}|}|$	$=$	$\overline{\overline{\overline{\;\;}|}|\,\overline{\;\;}|}|$	Condensation
	$=$	$\overline{\overline{\;\;}|}|$	Cancellation
	$=$		Cancellation

Thus, this expression has the value of the unmarked state.

Note that, in simplifying these expressions, we begin at the deepest space of the expression where there are no crosses which do not cross any other mark and apply the axioms of cancellation and condensation, and proceed 'outward' from there. There is, however, an alternative manner in which one can identify the value of such an expression. As one would expect in a calculus in which the values have dual static and dynamic meanings, this alternative interpretation reflects the dynamic nature of the expression. Here, one views the deepest space as transmitting signals up through the expression to be combined into a global valuation. The signal will change its state as it crosses from the inside of a distinction/indication to its outside.

An example, after Varela[13] will serve to illustrate the point. Let m stand for the marked state and let n stand for the unmarked state. Thus, mm = m; mn = nm = m; $\overline{\text{m}}|$ = n; and $\overline{\text{n}}|$ = m. Interpreting the form $\overline{\overline{\overline{\;\;}|\,\overline{\;\;}|}|\,\overline{\;\;}|}|$ in this manner we have:



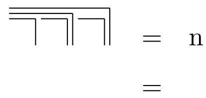
Transmitting

Transmitting

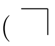
Condensation

Condensation
and transmitting

Thus:

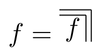


Again, this dual nature of the expression as an (dynamic) operator and (static) operand is central to the value of the calculus of indications. It is the key idea which makes the apparent paradoxes of self-reference which are inate in autonomous systems tenable.

What has been discussed so far is simply the arithmetic of the calculus of indications, since we have been solely concerned with the values of the calculus (in the same sense that $8 + 3 + 1 = 12$ deals with the arithmetic of natural numbers). Spencer-Brown does proceed to establish a complete consistent algebra of indicational forms which allows us to reason, if you will, about expressions involving variables, i.e. to establish the validity of forms comprising variables, regardless of the value of the variables ( or). While such an algebra is essential, it is not what is of primary importance here. Rather, we are interested in how to use the calculus of indications to describe systems which exhibit autonomy. However, the calculus as it has been presented so far is not adequate adequate for this purpose. Autonomous systems, and specifically autopoietic systems, exhibit a self-reference (i.e. an infinite recursive self-specification). That is, an autopoietic system enters into its own specification. In indicational terms, re-entry can be embodied by a statement that says a form, f say, is identical with a part of its contents. For example,

$$f = \phi(f)$$

where ϕ(f) is some indicational expression containing f as a variable. For example:



Varela adopts a convention which shows the point at which the form re-enters its own indicational space by extending the cross which contains the whole expression. Thus, $f = \overline{\overline{f}|}|$ is written

$$f = \overline{\overline{\sqcup}|}|$$

As one would expect, self-referential re-entrant expressions can be paradoxical. Consider the following:

$$f = \overline{f}|$$

Let us informally investigate this expression. Letting $f = \overline{}|$ and substituting in $f = \overline{f}|$, we have

$$\begin{aligned} \overline{}| \quad &= \quad \overline{\overline{}|}| \\ &= \end{aligned}$$

which is inconsistent: a marked state cannot be indentical to an unmarked state.

Similarly, letting $f = \quad$ and substituting in $f = \overline{f}|$, we have

$$= \quad \overline{}|$$

Again, the same paradox. The form $f = \overline{f}|$ is not demonstrable in the primary algebra since there is no arithmetic which will satisfy it. In Varela's words 'If f is marked then it is unmarked; if f is unmarked then it is marked'. This paradoxical or oscillatory nature of some re-entrant forms turns out to be a key to understanding autonomous systems.

As we have already mentioned several times, an indicational form has a dual interpretation as a static value (operand) or as a dynamic operator. Writing $f = \overline{f}|$ as a re-entrant form, we have

$$f = \overline{\sqcup}|$$

This is a value which is not reducible to a marked or an unmarked state.[1] However, the dynamic interpretation is even more intuitively pleasing. Adopting the previous technique of transmitting a signal from the deepest space up toward the cross containing the complete form

$$\begin{aligned} f \quad &= \quad \overline{f}| \\ &= \quad \overline{\sqcup}| \\ &\rightarrow \quad \overline{\sqcup}|\ \mathrm{m} && \text{Transmitting} \end{aligned}$$

[1]This, as Varela notes, is somewhat akin to the introduction of $i = \sqrt{-1}$ as an imaginary number in complex number theory.

That is, the form is marked. Making the re-entry explicit by re-placing the value of the form ($f = m$) back into the point at which it re-enters itself (effectively invoking $f = \overline{f}|$), we have:

$$\begin{array}{llll} f & \rightarrow & \boxed{m} & \\ & \rightarrow & \boxed{\ }\ n & \qquad \text{Transmitting} \end{array}$$

That is, the form is unmarked. Again, re-placing the value of the form ($f = n$) back into the point of re-entrance, we have:

$$\begin{array}{llll} f & \rightarrow & \boxed{n} & \\ & \rightarrow & \boxed{\ }\ m & \qquad \text{Transmitting} \end{array}$$

That is, the form is now marked. Thus the form $\boxed{\ }$ can be interpreted in a dynamic way as a temporal oscillation between the marked and the unmarked states:

$$\boxed{\ } = \rceil,\ \ ,\rceil,\ \ ,\rceil,\ \ ,\rceil,\ \ ,\ldots$$

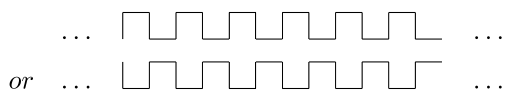
Note, however, that we arbitrarily chose to begin the process without reference to the 'current' state of $\boxed{\ }$ (which is either $\rceil$ or) by 're-placing' an unmarked state at the point of re-entry. Equally, we could have chosen to 're-place' a cross, in which case the sequence generated would be:

$$\boxed{\ } = \ \ ,\rceil,\ \ ,\rceil,\ \ ,\rceil,\ \ ,\rceil,\ldots$$

Thus,

$$\begin{array}{lll} \boxed{\ } & = & \rceil,\ \ ,\rceil,\ \ ,\rceil,\ \ ,\rceil,\ \ ,\ldots \\ & \textit{or} & \ \ ,\rceil,\ \ ,\rceil,\ \ ,\rceil,\ \ ,\rceil,\ldots \end{array}$$

This can be represented more graphically as



Both are valid interpretations of the form; the choice is made, *in context*, when we choose to begin to evaluate the sequence, i.e. at the time at which it has a marked or unmarked state. However, the two interpretations are mutually exclusive.

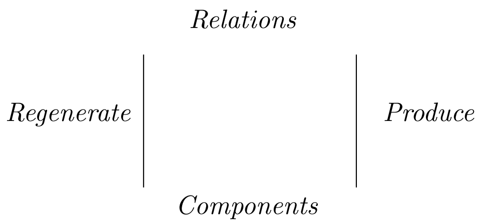
The introduction of this form is due to Varela and he extends the calculus of indications by developing the arithmetic to include $\boxed{\ }$ as a third explicit state, called the autonomous state (the state $\boxed{}$ arises autonomously by self-indication), and by constructing its attendant algebra. Varela then proceeds to address the dual dynamic interpretation of $\boxed{}$ (as two mutually exclusive waveforms) and develops a *waveform arithmetic* with an attendent algebra (which he calls a Brownian Algebra) for discussing these entities. As you would expect, waveforms of different form, but always comprising elements of marked and unmarked states, are generated by re-entrant forms of different types.

1.3 Relating Autopoiesis and the Calculus of Indications

To recapitulate, an autopoietic system is organised as a network of processes of production of components that:

- by interaction and transformation, continuously regenerate and realize the network that produced them;
- constitute it (the machine/system) as a concrete unity.

We can represent this situation diagramatically, as follows:



An understanding of this can be fraught with difficulty. First of all, it is very easy to misinterpret the definition $\left|\begin{matrix}R\\C\end{matrix}\right|$ in a simplistic and mechanistic manner, i.e., as some recursive computation. The $\left|\begin{matrix}R\\C\end{matrix}\right|$ pair is not a simple complementary couple co-existing at the one level (of abstraction). Rather the components C are *static* entities and the relations are *dynamic* entities; they are not constituents of the same domain. Thus, the relationship is not a complementary duality, $R \rightleftharpoons C$, but rather a higher order relationship. Varela makes the concept of a triadic relationship — a trinity — explicit by identifying a star operator which maps dualities to trinities. A star ($*$) statement is identically: 'the entity/the process leading to the entity'. We thus consider the entity and the process leading to the entity. The first essential point to be observed about such a statement (or operator) is that a trinity is formed if the two components mutually specify each other. For example, the relationship between a network and the trees constituting the network is that of a star:

$$* = \text{network/trees constituting the network}$$

In the case of autopoiesis, we have $* = R/C$, i.e., a triadic form comprising relations, components, and the mutual specifications across levels of abstraction.

In the remainder of this paper, we will adopt three equivalent notations to represent the concepts of relations and components: the indicational calculus, the star, and conventional set notation. Thus, if c_1 and c_2 are components, then a relation R involving c_1 and c_2 can be written:

$$\begin{aligned} R &= \overline{c_1\, c_2}\rceil \\ &= R/c_1 c_2 \\ R &= \{c_1, c_2\} \end{aligned}$$

However, the semantics of these three relations are not identical since the star relation involves a mutual specification of relation and component and it embodies a re-entrant definition and thus:

$$\begin{aligned} & * = R/c_1 c_2 \\ \equiv\ & R = \overline{R\, c_1\, c_2}\rceil \\ \equiv\ & R = \{c_1, c_2, R\} \end{aligned}$$

That is

$$\begin{aligned} & * = R/c_1 c_2 \\ \equiv\ & R = \lfloor\overline{c_1\, c_2}\rceil \\ \equiv\ & R = \{c_1, c_2, R\} \end{aligned}$$

Note that in the final set expression, the re-entrancy/recursion is implicit. However, as we will shortly see, it is a useful notation since it allows us to enumerate the possible reentrant relations for a given set of components and relations.

It is now possible to establish a link between the indicational calculus and the definition of autopoiesis. $\overline{}\rceil$ can represent both value and operator. Thus, an indicational expression can be viewed either as an organisation/structure of components or as a punctual embodiment of a dynamical relationship. This, it seems, is the value of the indicational calculus: it expresses the static and the dynamic aspects of autopoiesis in a single domain. That is, its rich semantics transcends both levels of autopoietic descriptions. This is particularly obvious in the case of the re-entrant value/expression $\lfloor\overline{}\rceil$ which is at once a collapsing of an infinite series of oscillating values and an embodiment of a self-specified self-referential system.

1.4 Realization of Autopoietic Systems

To make progress in studying autopoietic systems, it is desirable, at some stage, to attempt to synthesise such a system. This allows us to validate the theory, to identify its deficiencies and, hopefully, to correct them.

Autopoietic systems can, of course, be realised in may ways; that is inherent in the nature of organisation. Most familiar autopoietic systems are biological life-forms, but there are strong arguments that, for example, social entities can display autopoietic organisation. The interest here lies in the realisation of autopoietic systems in a computational domain, i.e., the objective is to generate computer-based autopoietic systems. The essential question, then, is: *What are the requirements of such a realisation of an autopoietic system?* The first issue we will address is the definition of the domain of discourse of the system, i.e. what are to be the components of the system, following which we will look at the structural requirements of realizing this system.

1.5 Structural Requirements of Realisation of Autopoietic Organisation

1.5.1 Definition of the Domain of Discourse

We must first identify the 'components of the universe of discourse', i.e., the elements of the stucture realising the autopoietic organisation, which are involved in the star relationship $* = r/c_i$.

We have two options: we can either invent a synthetic universe or we can attempt to abstract/identify what are the components of the real world.

It should be clear after our previous discussion in that the latter option is not plausible. Consequently, we adopt instead a minimalist approach in our simulation and choose elements of potential energy as the components of our universe. Thus, one assumes nothing but the possibility that quanta of energy can exist or be actualised and we postulate a (simulated) universe of energy. There is, in effect, only energy. Qualitatively, however, one may conceive of two types of energy: potential energy and actualised energy. Thus, we postulate that the universe is a universe of components of potential energy which have the capacity or potential to be actualised, i.e. to be 'pulled out of potential into existence'.

We will refer to the components of potential and actualised energy *pe* and *ae* respectively. Note that there is no quantitative difference between *pe* and *ae*; they are merely two perspectives of the same entity. In a sense, the transition or labelling is one which is meaningful only for an observer, i.e. an autopoietic system. We can now make some observations on the realisation of autopoietic systems in such a domain of discourse.

- Relations form between *pe* components, thus actualizing them ($pe \longrightarrow ae$).

- All relations are possible at any 'instant', i.e., all relations are always possible; the potential of any relation always exists.
- A relation is equally a component in any relation (itself or others).
- Any relation can form between any component or set of components.
- Autopoiesis implies and requires closure which, in turn, implies and requires re-entrance of the form characterising the autopoietic system (i.e. the characteristic form in re-entrant). Hence, the relations that bind the components can be re-entrant.

Relating this to the language of the indicational form which we will use to express the relations which will define our systems, we can say that if the 'depth' of the re-entrant form is minimal, i.e., it is a simple form, then the constituents of the re-entrant indicational expression (or form) are simply and necessarily components of the universe. That is,

$$f = \overline{f\,}\!\rceil \Longleftrightarrow f \in \{pe, ae\}$$

However, there is no requirement to limit the relations in this way: the depth can be greater than one and, hence, we can satisfy the requirement of re-entrance by allowing that the form (re-)enters it own indicational space at any level. Furthermore, the indicational sub-expressions or forms are equivalently relations of the components or, since every expression/form has a value, components themselves. Thus, the universe in which the autopoietic organisation is manifested/realised includes among its set of components not just $\{pe, ae\}$ but also the set of all possible relations. In this way, the possibilities for realisation of autopoietic organisation is now much richer than would first appear and one can now accommodate self-reference at any level. Additionally, the key concept of potentiality is preserved since relations are all potentially existing. The immediate ramifications of this is that, just as with 'elementary components', all (higher order) relations are possible at an 'instant' and any (higher order) relation can form between any other relation at any level, i.e., a relation can feedback into its own specification and re-enter its own definition.

Any structure which realises autopoietic organisation must embody the following elements.

1. A set of *pe* components, U, which is the universal set. As re-entrancy is allowed, relations can also be considered components and thus the universal set mut include these also.
2. A set of *pe* components, S, which can be perturbed (i.e. S is the sensor interface).
3. A set of *pe* components, M, which are capable of perturbing the environment (i.e. M is the motor interface).

4. A set of *pe* components, A, which comprises the remainder of the anatomy of the system.

   Thus, we have the following.

$$\begin{array}{l} S \cup M \cup A \subset U \\ S \neq \emptyset \\ M \neq \emptyset \\ S \cap A \text{ may or may not be} = \emptyset \\ S \cap M \text{ may or may not be} = \emptyset \\ M \cap A \text{ may or may not be} = \emptyset \end{array}$$

   Let $C = S \cup M \cup A$.

5. A structure to facilitate all possible (non-reentrant) relations, R, defined for this autopoietic system among all *pe* components, $pe \in S \cup M \cup A$, i.e., $pe \in C$.

6. A structure to facilitate all possible relations defined for this autopoietic system among all *pe* components, $pe \in S \cup M \cup A \cup R$. That is, a structure which facilitates re-entrant relations, since a relation formed of *pe* components and other relations, including itself, is identically a *pe* component.

7. The concept that a relation is actualised/operative must be embodied in some sense.

Requirements numbers 5 and 6 require a much more detailed discussion; this now follows. Requirement number 7 is discussed later in the section on implementation issues.

1.5.2 A structure to facilitate non re-entrant relations.

To recap, we are interested in a structure to facilitate all possible relations defined for this autopoietic system among all *pe* components, $pe \in S \cup M \cup A$, i.e.,$pe \in C$. The relations which are possible are given by the power set of $C, \mathcal{P}(C)$, and the number of possible relations is thus $2^{n(C)} = 2^c$, where $n(C) = c$ is the number of elements in the set C.

For example, the following relations are possible in a universe with three components, i.e. given $C = \{c_1, c_2, c_3\}$; $c = 3$,

$$\mathcal{P}(C) = \{\{c_1, c_2, c_3\}, \{c_1, c_2\}, \{c_1, c_3\}, \{c_2, c_3\}, \{c_1\}, \{c_2\}, \{c_3\}, \emptyset\}$$

Note, however, that not all of these relations are simultaneously possible since it is required that a component of a relation can only be involved in, at most, one relation. *The basis for this conjecture is that a component, which is either potential or actualised, is actualised by a relation (this is the basic axiom of autopoietic organisation) and it is only meaningful for it to be actualised by a* *single relation at any one instant.* We will refer to this conjecture as the *axiom of unique actualization.* This constraint necessitates that each of the relations are disjoint sets within the powerset of C. Let $R = \mathcal{P}(C)$, then we require the set D of possible non-reentrant relations, $D \subset R$, which is given by

$$D = \{R_i \mid R_i \in R \wedge R_i \cap R_j = \emptyset, \forall i, j; i \neq j\}$$

For example, given C (where $c = 6$, say), one possible set of relations is:

$$D = \{\{c_1, c_2, c_3\}, \{c_5, c_6\} \{c_4\}\}$$

The question arises as to the number of possible relations. Since $n(R) = r = 2^c$ it follows that $n(D) = d < 2^c$. However, the actual number depends of the choice of $R_1, R_2, R_3 ... R_i$. The constraint on the choice of R_i and R_j is that $R_i \cap R_j = \emptyset, i \neq j$. Thus, choosing R_i implies that R_j is a subset of the set of components C less the components included in R_i and hence R_j is an element of the powerset of this reduced set of components, *viz*:

$$\begin{aligned} \text{Choosing} \quad R_i &\Rightarrow R_j \subset C \setminus R_i \\ &\Rightarrow R_j \in \mathcal{P}(C \setminus R_i) \end{aligned}$$

Thus, we can simply choose an element of the powerset of C, i.e., choose a relation, and then the next relation which is possible is drawn from the powerset of $C \setminus R_i$. Now, we must again choose R_j since we are guaranteed that the elements of $\mathcal{P}(C \setminus R_j)$ are not disjoint. The next relation is, say, R_k, defined as follows.

$$R_k \in \mathcal{P}(C \setminus (R_i \cup R_j))$$

In general:

$$\begin{aligned} D &= \{R_i \mid R_i \in \mathcal{P}(C \setminus (R_{i-1} \cup R_{i-2} \cup \ldots))\} \\ &= \{R_i \mid R_i \in \mathcal{P}(C \setminus \cup (R_{i-j})), 1 \leq j \leq i, i \geq 1, R_0 = \emptyset\} \end{aligned}$$

1.5.3 A structure to facilitate re-entrant relations.

We now turn our attention to the extension of this development to include situations where relations of components and relations can form and, in particular, to the situation where the relation is defined in terms of itself, i.e., it is a re-entrant relation.

Again, $C = S \cup M \cup A$. In the previous case where we considered relations which were non re-entrant, we showed that the set of possible relations is given by:

$$D = \{R_i \mid R_i \in \mathcal{P}(C \setminus \cup (R_{i-j})), 1 \leq j \leq i, i \geq 1, R_0 = \emptyset\}$$

In this case, we must extend the set of components $\{pe_i\}$ from which relations can form to include the set of relations themselves. Thus:

$$D = \{R_i \mid R_i \in \mathcal{P}(U \setminus \cup (R_{i-j})), 1 \leq j \leq i, i \geq 1, R_0 = \emptyset\}$$

where U, the universal set, is given by:

$$U = C \cup D$$

Hence,

$$D = \{R_i \mid R_i \in \mathcal{P}\left((C \cup D) \setminus \cup (R_{i-j})\right), 1 \leq j \leq i, i \geq 1, R_0 = \emptyset\}$$

Note that this definition of D is recursive. Thus, the number of elements in U is infinite ($u = n(U) = \infty$) since a relation of components (i.e. elements of U) will form a new component, increasing the value of u by one and, in the process, producing (at least) one new relation, which is then an element of U, again increasing u by one, and so on *ad infinitum.* The problem is to deal with this recursion in a coherent and meaningful manner.

To overcome this difficulty, we will introduce the the concept of a degree of a relation defined to be the level of recursion to which it is necessary to go to generate the relation. For example, a relation of *pe*s ($pe \in C$) needs just one instantiation of the definition of D and is of degree 1; a relation of relations of degree 1 (and, perhaps, $pe \in C$) would require two invocation of this recursive definition and is, hence, of degree 2. Thus we can successively re-write the definition of D:

$$\begin{array}{lcl}
D^1 & = & \left\{R_i^1 \mid R_i^1 \in \mathcal{P}\left(C \setminus \cup \left(R_{i-j}^1\right)\right), 1 \leq j \leq i, i \geq 1, R_0^1 = \emptyset\right\} \\
D^2 & = & \left\{R_i^2 \mid R_i^2 \in \mathcal{P}\left((C \cup D^1) \setminus \cup \left(R_{i-j}^2\right)\right), 1 \leq j \leq i, i \geq 1, R_0^2 = \emptyset\right\} \\
D^3 & = & \left\{R_i^3 \mid R_i^3 \in \mathcal{P}\left((C \cup D^1 \cup D^2) \setminus \cup \left(R_{i-j}^3\right)\right), 1 \leq j \leq i, i \geq 1, R_0^3 = \emptyset\right\} \\
\vdots & & \vdots \\
D^n & = & \left\{R_i^n \mid R_i^n \in \mathcal{P}\left((C \cup D^1 \cup D^2 \ldots \cup D^{n-1}) \setminus \cup \left(R_{i-j}^n\right)\right), 1 \leq j \leq i, i \geq 1, R_0^n = \emptyset\right\}
\end{array}$$

The trailing superscript on D and R denotes the degree. If we define D^0 to be C, then we can write:

$$D^n = \left\{R_i^n \mid R_i^n \in \mathcal{P}\left(\cup \left(D^k\right) \setminus \cup (R_{i-j^n})\right), 1 \leq j \leq i, i \geq 1, 0 \leq k \leq n-1, R_0^n = \emptyset\right\}$$

If the relation is re-entrant, then $d = \infty$. This expression embodies the restriction/axiom which we invoked stating that a component of a relation can only be involved in, at most, one relation, i.e. components of relations are mutually exclusive. This effectively means that, in the case of non re-entrant relations, the structure which is required is a tree-like. Consequently, any given variable can only appear once in an expression describing autopoietic organisation in the calculus of indications.

Before we can identify a structure which will facilitate the implementation of this recursive definition, we need to see whether or not there are any other restrictions, arising from the extension of the system to re-entrant relations, which we can place on the formation of relations and their use in other relations.

While re-entrant relations must still respect the axiom of unique actualization (or a component or relation: i.e. a *pe* in general), and thus a tree structure is still required, it does allow for the relation to re-enter at any point in the sub-tree of which it is a root. Thus, the expression can be represented as a non-reticulated tree, and does not required a general graph representation: i.e. graphs are invalid. This provides us with a solution to the problem of identification of the required structure for the realisation of autopoietic systems: a connection between a relation node and each of the relation nodes in its sub-tree is required in general.

Thus, the required stucture is a threaded tree of relations of arbitrary degree, such that each thread only connects a root with a node in its sub-tree, and where each element of the set D^n

$$D^n = \left\{ R_i^n \mid R_i^n \in \mathcal{P}\left(\cup\left(D^k\right) \setminus \cup\left(R_{i-j}^n\right)\right), 1 \leq j \leq i, i \geq 1, 0 \leq k \leq n-1, R_0^n = \emptyset \right\}$$

is represented in the tree. However, since all relations are potentially possible, i.e. the threaded tree is the 'instantaneous' actualizated structure which is defined by the 'choice' of R_{i-j}^n in the expression above, the system must be capable of instantiating all of these threaded trees and, consequently, the required 'super-structure' is a fully-connected graph.

At this point, we have a complete characterisation of the type of structure that is required for the realisation of an autopoietic system. The goal now is to implement it, i.e., to develop a simulator for autopoietic systems. We will proceed to complete this section on the realization of autopoietic systems by describing the implemenation of a simulator. We will then proceed to discuss the development of an autopoietic system, *per se*.

1.6 The Development Environment: An Autopoeitic Simulator

We will begin this discussion by identifying the components of the computer architecture which supports the autopoietic simulator and then we will proceed to discuss the software environment within which the simulation is achieved.

We noted previously that the autopoietic system comprises three sets of potential elements, S, M, and A, corresponding to elements which form the sensor surface (i.e. which are perturbable by the environment) those which form the motor surface (i.e. which are capable of perturbing the environment), and the remainder of the anatomy. This implies that the system must comprise a sensor surface, a motoric mechanism, and a sub-system capable of supporting the actualisation of potential elements. In this simulator, we have chosen to adopt a CCD TV camera as the (light sensitive) sensor surface and a robot which holds the camera. Thus the autopoietic system, as embodied in the sensor surface, can move in the environment by appropriate actualisation/actuation of the motor surface, i.e. by moving the robot. A PC-hosted transputer microprocessor completes the inventory of physical devices which fulfill each of the three rôles (see figure 1).

Figure 1: Architecture of the Autopoietic Simulator.

1.6.1 The Camera System.

The camera system comprises a conventional Sony CCD monochrome sensor which generates a CCIR video signal. The amplitude of this signal is dependent on the intensity of the light incident on the sensor. The camera is connected to a framestore which digitises the analogue CCIR video signal, sampling and quantising it to produce an array of 512x512 8-bit values representing the intensity of light incident on the CCD sensor. The framestore resides in an Olivetti M380 PC where the image can be accessed by the transputer board which supports the software environment to be described in the second next section. The image is accessed by the transputer system via a program, normally referred to as the Alien File Server (AFS), which runs on the PC.

1.6.2 The Robot System.

Movement of the autopoietic system is accomplished by mounting the sensor set on the hand of a 5° of freedom manipulator. Although the remaining two sets of potential elements, M and A, are physically elsewhere (and static), this makes no difference from a logical point of view: if the set S of potential elements by which the system is perturbed by the 'world' is free to move as a result of perturbation of the motor surface, this is equivalent to a freely moving autopoietic or allopoietic system. Although the robot has five degrees of freedom, it only uses three of these: rotation of the base, shoulder, and elbow (see figure 1). This allows the system to position the sensor/camera anywhere in the working envelope of the robot, but the orientation of the camera is not controlled and is fixed in the hand, pointing straight down along a vertical line of sight. Since the base motor is not coupled to the wrist motor, a rotation of the base will cause a significant change in the field of view of the camera, i.e. the movement of the sensor surface (which is directionally sensitive to light) is not purely translational but also comprises a rotational component. Since the distance of the camera to the axis of rotation in the base is large (30 cm), the rotational effect is small *w.r.t.* the translational effect. The robot is controlled by the transputer system via a RS232 serial link between the controller board and the PC.

1.6.3 The Transputer-based Computer System.

The PC hosts a transputer board, comprising a T800 transputer and 2 Mbytes of memory, on which the simulation program runs. The transputer communicates with the framegrabber (to acquire the images representing the current state of actualisation of the sensor potential elements) and the robot (which accomplises actualisation of the motor surface) via the AFS which runs on the PC CPU. The transputer runs a complete software environment call TDS, for Transputer Development System, and is programmed in *occam*, a language which supports parallel programming and, in particular, which facilitates the execution of multiple concurrent asynchronous processes. This is obviously of paramount importance since the processes involved in instantiating autopoietic systems, i.e. the actualisation of potential elements and their relations, are asynchronous and concurrent. This is an issue to which we will return later in the section on implementation issues where will will describe in detail the occam language and the manner in which potential elements and relations are simulated.

1.6.4 Implementation Issues for the Realisation of Autopoietic Structure

A realisation of an autopoietic system immediately implies the realistion of two things: components and relations. These components and relations must exist concurrently and all processes must proceed concurrently. Note that true concurrency and autonomy of processes is required. This is not the same as requiring *parallelism* since parallelism tacitly involves the idea of doing things 'at the same time'; while there might be no causal interconnnection between parallel processes, there is the implicit understanding that they are synchronised at certain critical points. They, typically, have coordinated beginnings and ends. Since it is intended to realise these systems in a computational domain (using the occam language), it is necessary to first address the correspondence between occam entities, and relations and components. Occam itself comprises two essential building blocks: processes and channels. Processes communicate via channels and each channel provides a one-way connection between two concurrent processes. Occam programs are constructed from three primitive processes: assignment (:=) which changes the value of a variable; input (?) which allows a value to be received from a channel; and output (!) which allows a value to be sent to a channel. Processes are combined to form sequential, parallel, or alternative constructs using the keywords SEQ, PAR, ALT, respectively.

Since channels are use for linking, or relating, processes, it would be natural to associate channels with relations and processes with components when realising the system. However, it is interesting to note that one can also interpret the concepts in an alternative manner such that the component is the channel and the relation is the process. This is appealing for two reasons. Firstly, relations and transformations are intrinsically dynamic active conceptual entities; this is well modelled by the active nature of the occam process. Secondly, components are means to a realisation of the network and are intrinsically concrete and passive (they're there); this is well modelled by a communication channel.

A potential element *pe*, then, is implemented as a channel. The type of the channel is defined as a *component.protocol*, *viz:*

```
PROTOCOL component.protocol – define the type of communication
CASE                        – which will take place over a channel:
  active                    – channel sends "active" when component exists
:                           – i.e. when PE is actualised
```

Thus, the *pe* is actualised instantaneously when the value active is sent on a channel. Recurrent actualisation requires recurrent sending of the active value. In this way, the actualisation of the potential element, over time, is manifested as a pulse train of temporally-varing frequency. The three classes of components, i.e. sensor surface, action surface, all other potential elements, are implemented as an array of channels, defined as follows.

```
VAL INT number.of.relations IS 30

– instantiate two 1-D array of channels to allow sensory input to
– and action output from each relation

[number.of.relations]CHAN OF component.protocol sensory.input, action :

– instantiate a 2-D array of channels with n x n channels in total to allow
– each of n relations to communicate

[number.of.relations][number.of.relations]CHAN OF component.protocol component :
```

A relation is implemented as a procedure, i.e. a named process, which relates *sensory.input*, *action*, and *components*, *viz*:

```
PROC relation (CHAN OF component.protocol sensory.input,
                                          monitor.output,
                                          action,
  [number.of.relations][number.of.relations] CHAN OF
  component.prototcol component,
```

VAL INT relation.number)

Thus, each relation forms a relationship between a single sensory potential element (component), a single action potential element, and as many relation potential elements as there are relations. An additional *monitor.output* potential element is also included as an aid to assist in debugging the system. Hence, if there are n relations, there must be n^2 components since each relation is potentially a component of each other relations. If the relation (which is, after all, a potential element) is actualised, this is signalled by activating each (output) component channel, effectively communicating with each other relations. Unless precautions are taken, such concurrent communication among a large network of processes (i.e. relations) would inevitably lead to deadlock, wherein processes waiting to send communication messages are kept waiting indefinitely because the destination process is simultaneously waiting to send a message back. Deadlock is avoided in relation processes by buffering all input and, thus, decoupling the outgoing and the incoming communications. This is perhaps best described visually, as depicted in the figure 2. One essential point to notice about this organisation is that each input is buffered via a *single* status buffer with a value — active or inactive — which is continually read by the process handling the output and continually overwritten by the process handling the input. It is thus possible that information, i.e. discrete actualisation can be 'lost'. However, this does not adversely affect the performance of the system since such losses are effectively 'integrated out' over a short period of time.

Let us now recall the two fundamental premises of autopoietic organisation: (a) that the components, by interaction and transformation, regenerate the (network of) relations and (b) that the network of processes, i.e., the relations produced by the components. Is this feasible with the schema described above? Firstly, the relations do produce the components, in the sense that the processes certainly produce 'links' or channels by communicating on them; all channels exist in potentiality and are 'activated' or produced by the process. Secondly, the components do produce the relations in that all processes exist in potientiality and are activated by communication on a channel.

1.7 Initial Experiments and System Verification

To verify that the system operates correctly, i.e. to verify that potential elements can be actualised concurrently and that communication takes place without deadlock, a simple allopoietic (i.e. externally-governed) system for tracking targets was configured.

Four receptive fields on the sensor surface were defined as shown in figure 3. These four symmetrically organised receptive fields can be parameterised by r, the radius of the receptive field, and d, the separation of two symmetrically

Figure 2: Internal organisation of a relation process.

Figure 3: Spatial arrangement of receptive fields constituting S.

opposed fields. If d is appropriately tuned to the size of the object to be tracked (d is less than the width and the height of the object), the allopoietic organisation shown in the following figure will tend to cause the robot/sensor to centre itself over a given target (object) which exhibits homogeneous reflectance. As the object moves, the system will react to keep the four receptive fields within the homogeneous region, i.e. it will track the object. In conventional parlance, the system tracks homogeneous regions of a particular size.

We can characterise this system by the following set of productions:

$$\begin{array}{ccc} s_i & \longrightarrow & r_i \\ r_i & \longrightarrow & a_i \end{array}$$

where $i = 1..4$.

In terms of the Calculus of Indications, we have the following trivial relations

$$\begin{array}{ccc} s_i & = & r_i \\ r_i & = & a_i \end{array}$$

The structure of this allopoietic system is shown diagramatically in figure 4.

Since we are dealing with an allopoietic system, i.e. a system which 'blindly' responds to external changes in order to achieve some predetermined goal (maintenance of homogeneous regions within the centre of the receptive fields) it might be instructive to try to classify such behaviour according to conventional control theory. In classical control theory, control can be effected by three (super-imposable) modules or controllers:

- a proportional controller, whereby the reaction of the system is proportional to the difference between the required position (termed the set point, in conventional parlance) and the current position;

Figure 4: Structure of allopoietic system; version 1: proportional control.

- a differential (or derivative) controller, whereby the reaction is a function of the temporal rate of change between the required position and the current position.
- an integrative controller whereby the reaction is a function of the accumulated (or average) discrepancy between the required position and the current position.

Given the productions $s_i \longrightarrow R_i$ and $R_i \longrightarrow a_i$ above, it is clear that this allopoietic system is exhibiting proportional control, that is, the degree of activity is proportional to the signal received at the sensor surface. If we wish to incorporate differential control, we must establish some form of temporal differentiation of the signal. We can do this by adding the following production:

$$s_i \wedge \overline{R_i} \longrightarrow R_i$$

where the symbol $\wedge$ denotes logical AND and the overbar denotes logical negation. In essence, this production makes the relation R_i re-entrant and incorporates the logical negation of the state of actualisation of the relation, in this case at the previous instant. Thus, this production means that the relation is actualised when the relation is currently actualised and if it wasn't previously actualised. Actualisation arises when there is a (temporal) difference between subsequent states of actualisation. In order to represent this production using the Calculus of Indications, we need first to note that the logical conjunction AND of two Boolean variables $A \wedge B$ has an equivalent form in the Calculus, *viz:*

$$A \wedge B = \overline{\overline{A|}\,\overline{B|}\,}|$$

which can be clearly seen from table 1 where the marked state is equivalent to the value TRUE, the unmarked state as the value FALSE, and logical negation is effected by the cross operator. Thus, the production

$$s_i \wedge \overline{R_i} \longrightarrow R_i$$

can be expressed:

$$\begin{aligned} R_i &= \overline{\overline{s_i|}\,\overline{\overline{R_i|}|}\,}| \\ &= \overline{\overline{s_i|}\,R_i\,}| \\ &= \overline{\overline{s_i|}\,\overline{\;\;|}\,}| \end{aligned}$$

Combining this with the first production, $s_i \longrightarrow R_i$ (i.e. $s_i = R_i$), we have:

$$R_i = \overline{\overline{s_i|}\,\overline{\;\;|}\,}|\, s_i$$

The structure of this second allopoietic system is shown diagramatically in figure 5. The system as it is currently constituted runs reasonably quickly with approximately 200ms elapsing between dissemination of sensed data and the activity of the robot motors. The main delay arises because of the overhead in the manner in which channels effect actualisation, i.e., by communicating pulses. Thus, analogue values are effected by a form of Pulse Code Modulation (PCM) technique and, hence, a sensed light intensity of, say, 127 in a grey-scale range of 0-255 would require 127 distinct pulses to be communicated. Since the physical dynamics of the system is a function of the sensed light data, the system damping can be controlled by reducing the quantisation resolution of the sensor, either by software or by decreasing the aperture of the lens, thus decreasing the value associated with the light intensity.

2 Self-Renewing, Autopoietic, Systems

2.1 Realizing the Autopoietic System

The behaviour exhibited be the allopoietic system which was described in the last section is due solely tothe pre-wired structure of the system: it tracks homogeneous regions because it is built that way. It reacts to the changes in the visual environment in a pre-determined manner and, as such, is effectively controlled by the environment. It is truly an allopoietic system.

A	B	$\overline{A}\rvert$	$\overline{B}\rvert$	$\overline{A}\rvert\,\overline{B}\rvert$	$\overline{\overline{A}\rvert\,\overline{B}\rvert}\rvert$
		⌉	⌉	⌉	
	⌉	⌉		⌉	
⌉			⌉	⌉	
⌉	⌉				⌉

Table 1: The Boolean logical AND in the Calculus of Indications.

Figure 5: Structure of allopoietic system; version 1: proportional control.

An autopoietic system, on the other hand, will develop the structure itself and then maintain that structure. This emergence of structure is, of course, explicitly dependent on the class of relations which constitute the autopoietic system. This begs the essential question: What are the class of relations, i.e. productions, which are required to give rise to autopoietic organisation? This is the central issue to which we now turn.

The current working hypothesis is that the requisite relations are those which give rise to the non-reticulating hierarchical, but re-entrant, structure derived in the previous chapter. Thus, the productions must meet the following requirements.

1. The productions must enable the recurrence of the structure; the structure must be self-producing (self-organising).

2. The productions must facilitate the re-organisation arising from perturbations in $pe \in S$ and $pe \in M$ (i.e. the potential elements comprising the sensor surface and the motor surface).

3. The inclusion or exclusion of a pe in S, M, or A, must be a function of the structure itself and will be dynamic. Thus, the system is 'compositionally' open (with free transfer of potential elements into and out of the structure) but organisationally closed, i.e. it is a self-specifying closed self-referential system which maintains its identify in spite of the flux of constituents into and out of the structure.

In order to implement an autopoietic system, we must constrain the set of relations. We can do this specifically by placing restrictions on the degree of the tree (equivalently, the number of components whch can be involved in a relation); and the depth of the tree (equivalently, the depth of an indicational form or the number of relations that can emerge). These are necessary conditions for the realisation of autopoietic structures but it seems clear that they are not sufficient. The question remains, then, as to what are the set of necessary and sufficient conditions?

We turn to *Systematics* to help provide the answer. Our contention, and our working hypothesis, is that an autopoietic system is a quinquepotent system; i.e., it requires a minimum of five terms for its realization. The definition of an autopoietic system is clearly that of a aelf-renewing autonomous system but it falls short of self-reproducing, self-regulating, or self-directing autonomy. For these levels of existence, additional terms are required according to the spectrum of autonomy which we identified in the first part of the paper. In developing (a simulation of) a quinquepotent entity, each term in a system must have very specific attributes. We begin by developing our autopoietic system from a one term system, rising to five, as follows.

- *Number of terms:* One.
  Terms: Universe.

Category: Wholeness.

- *Number of terms:* Two.
  Terms: Universe; System.
  Category: Polarity; complementarity.
- *Number of terms:* Three.
  Terms: Universe; System; Interface.
  Category: Relatedness; dynamism.
- *Number of terms:* Four.
  Terms: Universe; System; Interface; Interaction.
  Category: Subsistence; activity.
- *Number of terms:* Five.
  Terms: Universe; System; Interface; Interaction; Development.
  Category: Potentiality; Significance.

The first four terms are quite easy to identify and the four term system belongs to the domain of controllable systems with which we are familiar. The transition from a four term system to a five term system is more significant since it represents a transition from the hyponomic to the autonomic level. It is this critical increase in complexity that facilitates the emergence of autonomy and we must be very careful to ensure that the fifth term is distinct from the other four, while at the same time contributing to their specification. Note that all of the first four terms (Universe, System; Interface; and Interaction) can be validly defined in space and time. In identifying the fifth term, it is necessary to go beyond this 4-D manifold and introduce the eternal aspect we discussed above.

We postulate that the fifth term then is 'Develoment in eternity' which, from our perspective, is actualization of potential and self-organization. In Bennett's relativistic ontology, this takes place outside time and space. If development is the term, then what is developed? We speculate that it is order: to promote the transition from potential to actual; i.e. *the increase the incidence of actualisation*. This means that an autopoietic system *never* rests in a state of equilibrium. Consequently, the behaviour which would (and is) observed in an autopoietic system is a 'restless' striving: always self-specifying, always self-defining; exactly the behaviour of an autonomous entity. Let it be clear that this fifth term is no *vital force*. It is not a force at all — it has no existential quality: it is 'merely' the inclusion of an additional systemic aspect through which the autonomy of the system arises. We will refer to this term as 'in-formation' (after Varela) since this name captures the essence of the term: formation and actualization of structure from within and it reflects the self-renewing aspect of quinquepotent entities. It remains to be demonstrated that it is possible to incorporate this new requirement in the development (or design) of an artificial autopoietic system.

Returning to Systematics, there are ten primary relationships in the five term system:

$$\begin{array}{lcl}
\text{Universe} & \longrightarrow & \text{System} \\
\text{Universe} & \longrightarrow & \text{Interface} \\
\text{Universe} & \longrightarrow & \text{Interaction} \\
\text{Universe} & \longrightarrow & \text{In} - \text{formation} \\
\text{System} & \longrightarrow & \text{Interface} \\
\text{System} & \longrightarrow & \text{Interaction} \\
\text{System} & \longrightarrow & \text{In} - \text{formation} \\
\text{Interface} & \longrightarrow & \text{Interaction} \\
\text{Interface} & \longrightarrow & \text{In} - \text{formation} \\
\text{Interaction} & \longrightarrow & \text{In} - \text{formation}
\end{array}$$

Each of these relationships plays a critical rôle in the specification of the system. One of particular interest here is the 'Interface/In-formation' relation since it both requires and allows the dynamism of the interface to provide the transition from *pe* to *ae*, resulting in a materially-open but organisationally-closed system. That is, the set of potential elements which constitute S, the sensor surface, is being dynamically re-defined by the sytem itself. It is possible that this applies equally to C, the set of components of the autopoietic system. How does this arise? Since S is an (unspecified) subset of U, if we include $s_i \in S$ in an autonomous form then it can enter into the form or not through its interaction with other potential elements, specifically $c_i \in C$ or, perhaps, $a_i \in A$.

Now consider f = $\overline{\overline{s_i|}\;\overline{a|}\,|}$. Here, f = $\overline{\quad|}$ when $s_i = \overline{\quad|}$ iff $a = \overline{\quad|}$. That is, f is actualised iff s_i and a are both marked: see table 2. In a sense, a conditions the equivalence of f and s_i. Now let $g = g(f)$, i.e., let f be a term in g. Now also let $a = g$, i.e. the form of g re-enters f (and g) as the conditioning mark, *viz.*:

$$g = \overline{\begin{array}{l}\vdots \\ \overline{\overline{s_i|}\;\underline{\overline{\quad|}|}\,|}\ldots\end{array}\Big|}$$

If g is marked, then the s_i will be incorporated into the form g; if g is autonomous, this implies that s_i will, all things being equal, be periodically an element of g.

Now let us proceed to consider a simple variant. Let $m_i = \underline{\overline{|s_i|}}$ and, in the following, let m stand for $\overline{\quad|}$, the marked state and let n stand for the unmarked

s_i	a	f
	$\urcorner$	
$\urcorner$		
$\urcorner$	$\urcorner$	$\urcorner$

Table 2: Conditioning of s_i by a.

state.

$$
\begin{array}{lcccl}
\text{Let } \mathrm{s_i = n} & \Rightarrow & m_i & = & \lfloor\overline{\mathrm{n}|} \\
 & & & = & \lfloor\overline{\mathrm{n}|}\,\mathrm{m} \\
 & & & = & \mathrm{n} \\
 & \Rightarrow & m_i & = & \lfloor\overline{\mathrm{m}|} \\
 & & & = & \lfloor\overline{\mathrm{m}|}\,\mathrm{n} \\
 & & & = & \mathrm{m} \\
\text{Thatis} & \Rightarrow & m_i & = & \lfloor\overline{\quad|} \\
 & & & & \\
\text{Let } \mathrm{s_i = m} & \Rightarrow & m_i & = & \lfloor\overline{\mathrm{m}|} \\
 & & & = & \lfloor\overline{\mathrm{m}|}\,\mathrm{n} \\
 & & & = & \mathrm{n} \\
 & \Rightarrow & m_i & = & \lfloor\overline{\mathrm{n\ m}|} \\
 & & & = & \lfloor\overline{\mathrm{m}|} \\
 & & & = & \lfloor\overline{\mathrm{m}|}\,\mathrm{n} \\
 & & & = & \mathrm{n} \\
\text{That is} & & m_i & = & \mathrm{n}
\end{array}
$$

Hence, if $s_i = \mathrm{n}$, which is the unmarked state, then m_i is an autonomous form; if $s_i = \mathrm{m}$, which is the marked state, then m_i is a stable form, in this case unmarked (see table 3).

In a similar manner, letting $m_i = \overline{\overline{s_i|}\lfloor\quad|}$ and we have the behaviour shown in the table 4.

We can view this second case in the following manner. A motor surface potential element, m_i, is a compensation for the external actualisation (perturbation) of the sensor potential element. If $s_i = \overline{\quad|}$, then the motoric component

s_i	$m_i = \overline{s_i}\rvert$
	$\overline{s_i}\rvert$
$\overline{\quad}\rvert$	

Table 3: Autonomously-conditioned variable.

s_i	$m_i = \overline{\overline{s_i}\rvert\ \overline{\quad}\rvert}$
$\overline{\quad}\rvert$	$\overline{s_i}\rvert$

Table 4: Another autonomously-conditioned variable.

is also not actualised; if $s_i =$ then m_i assumes an autonomous value, which is astable and, in a sense, is a form of indeterminacy.

Let it be clear that there are two distinct issues here:

1. The conditioning of s_i (or a_i or c_i) so that they are included or excluded from the form, i.e. we are concerned with the material openness of the system while maintaining its organisational closure. In this instance, it is the form which re-enters and conditions.

2. The conditioning of m_i by either the form f or, in particular, by s_i, i.e. the compensation by m_i for the external actualisation (perturbation) by s_i.

The systemic activity referred to in (1) provides the autonomous form required of autopoietic organisation, while that in (2) provides the compensation for perturbation to facilitate that in (1). Any true autopoietic system which can be perturbed must, therefore, include both of these aspects.

So now, finally, we can begin to formulate (a possible set of) necessary and sufficient conditions of autopoietic organisation in terms of the calculus of indications.

Recall that, viewed in the calculus of indications, an autonomous autopoietic form gives rise to periodic (waveforms) of marked and unmarked states. The transition in state is both cyclic and unstable (or, more accurately, self-stabilising, self-organising, and self-actualizing). I postulate that this is a necessary and sufficient condition for autopoietic systems. If this is so, and it remains to be proved, then it follows from the previous argument that a system is autopoietic *iff* it satisfies the indicational equation:

$$\begin{aligned} f &= \{\{\overline{\;\;\;}|\}\{\;\;\;\}\} \\ &= \{\{\mathrm{m}\}\{\mathrm{n}\}\} \end{aligned}$$

Incorporating (1) and (2) above, and exploiting the results of the previous discussion on the depth and degree of an expression, we can re-write f as follows.

$$f = \overline{\vdots\;\overline{\overline{m_j|}\,\overline{\;\;}|}|\ldots}|$$

but

$$m_j = \overline{\overline{s_i|}\;\;\;}|$$

Thus:

$$f = \overline{\vdots\;\overline{\overline{\overline{s_i|}\;\;\;}|\,\overline{\;\;}|}|\ldots}|$$

Thus, actualisation of a potential element m_j (and, equally, a_j) is dependent on s_i and the form f itself, such that the actualisation gives rise to the requisite cyclicity of self-stabilising actualisation.

We are now in a position to design our simple autopoietic system. Let us choose the simplest form of f, as given in the indicational equation above. That is, let f be defined as follows:

$$f = \overline{\overline{\overline{m_j|}\,\overline{\;\;}|}|}$$

where

$$m_j = \overline{\overline{s_i|}\;\;\;}|$$

Thus:

$$f = \overline{\overline{\overline{\overline{s_i|}\;\;\;}|\,\overline{\;\;}|}|}$$

And let $i, j = 1..4$ so that we have the same anatomy as with the allopoietic system. This form f can be captured by seven productions, as detailed in table 5.

Note that the production $P0 : s \longrightarrow s$ is trivial. Why then is it necessary? It concerns the problem of mapping sensory inputs into the system. Earlier, we defined the structure realizing an autopoietic form as a set $D = f((S \cup M \cup A \cup R)$. Thus, we allowed for a distinction between S (sensor components), M (motor components), and R (relations). We have a finite set of s_i components which defines the physical instantiation of the autopoietic system.

Label	*Production*	*Equivalently*	*form*
$P0$	$s_i \longrightarrow s_i$		s_i
$P1$	$s_i \longrightarrow \overline{s}$	$P0 \longrightarrow \overline{P0}$	$\overline{\ \ }\vert$
$P2$	$\overline{s_i} \vee m_j \longrightarrow \overline{\overline{s_i} \vee m_j}$	$P1 \vee P2 \longrightarrow \overline{P1 \vee P2}$	$m_j = \overline{\overline{s_i\vert}\ \sqcup\vert}$
$P3$	$m_j \longrightarrow \overline{m_j}$	$P2 \longrightarrow \overline{P2}$	$\overline{m_j\vert}$
$P4$	$f \longrightarrow \overline{f}$	$P6 \longrightarrow \overline{P6}$	$\overline{f\vert}$
$P5$	$\overline{m_j} \vee \overline{f} \longrightarrow \overline{\overline{m_j} \vee \overline{f}}$	$P3 \vee P4 \longrightarrow \overline{P3 \vee P4}$	$\overline{\overline{m_j\vert}\ \overline{f\vert}\,\vert}$
$P6 = f$	$\overline{\overline{m_j} \vee \overline{f}} \longrightarrow \overline{\overline{\overline{m_j} \vee \overline{f}}}$	$P5 \longrightarrow \overline{P5}$	$\overline{\overline{\overline{m_j\vert}\ \overline{\sqcup\vert}\,\vert}\,\vert}$

Table 5: Productions for the Autopoietic System.

Label	*Production*	*Equivalently*	*form*
$P0$	$s_i \longrightarrow s_i$	$s_i \longrightarrow P0$	s_i
$P1$	$s_i \wedge \overline{m_j} \longrightarrow P1$	$P0 \wedge \overline{P1} \longrightarrow P1$	$\overline{\overline{s_i\vert}\ \sqcup\vert}$
$P2$	$\overline{m_j} \vee \overline{f} \longrightarrow P2$	$\overline{P1} \vee \overline{P2} \longrightarrow P2$	$\overline{\overline{\overline{m_j\vert}\ \overline{\sqcup\vert}\,\vert}\,\vert}$

Table 6: Alternative Productions for the Autopoietic System.

Equally, we have a finite set of m_i components. To allow for the implementation (or instantiation) of D, we just need to allow for the possibility of any $s_i \in D_j$. There are two ways of doing this:

1. $s_i \longrightarrow D_j, \forall j$, i.e., by mapping s_i into every possible D_j.

2. $s_i \longrightarrow R_j, \forall j$, i.e., by mapping s_i into every possible initial relationship.

Option 2. is simpler and merely requires one additional production in the set of relations R, *viz.* $s_i \longrightarrow s_i$. Thus, a perturbation is a component which is potentially included in all other relations.

Having indentified the production and, identically, the relations, we must now implement them. There are two ways of doing this. On the one hand, we could simulate a (potential) population of relations for *each* production. Alternatively, we could simulate a (potential) population of generic relations embodying all possible productions but enforce mutually exclusive actualization of one of the relations/productions. That is, when actualized, a relation 'represents' only one production. There seems to be a problem with the implementation of these relations. It hinges upon what we say is a relation. Is it a relationship which can be non-trivial?, e.g., $\overline{\overline{c_1|}\,\overline{c_2|}\,|}$ or $\overline{c\,|\,\sqcup\,}|$, or must a relation be restricted to just one level crossing, *viz:* $\overline{c\,}|$ and hence we would have three relations in a form such as $\overline{\overline{c_1|}\,\overline{c_2|}\,|}$. This has a big influence on how we implement the system for we either have the seven productions detailed in table 5 or three, more complex, relations as follows.

1. We have $P0$ as before:
$$s_i \longrightarrow P0$$

2. $P1$ arises from the term: $m_j = \overline{\overline{s_i|}\,\sqcup\,}|$, that is:
$$\begin{aligned} m_j &= \overline{\overline{s_i|}\,\sqcup\,}| \\ &= \overline{\overline{s_i|}\,m_i}| \\ &= \overline{\overline{s_i|}\,\overline{\overline{m_i}|}|}| \end{aligned}$$
Thus, production $P1$ is:
$$s_i \wedge \overline{m_j} \longrightarrow P1$$
But $s_i = P0$ and $m_j = P1$, and hence production $P1$ becomes:
$$P0 \wedge \overline{P1} \longrightarrow P1$$

3. The final production arises from the expression $f = \overline{\overline{\overline{m_j}\rceil\,\overline{\underline{}}\rceil}\rceil}\rceil$. That is,

$$
\begin{aligned}
f &= \overline{\overline{\overline{m_j}\rceil\,\overline{\underline{}}\rceil}\rceil}\rceil \\
&= \overline{\overline{\overline{m_j}\rceil\,\overline{f}\rceil}\rceil}\rceil \\
&= \overline{m_j}\rceil\,\overline{f}\rceil
\end{aligned}
$$

Thus, production $P2$ is:

$$\overline{M_j} \vee \overline{f} \longrightarrow P2$$

Again, $m_j = P1$ and $f = P2$, and hence production $P2$ becomes:

$$\overline{P1} \vee \overline{P2} \longrightarrow P2$$

These three productions are summarized in Table 6. And so, which set of productions do we choose? Which is correct? To answer this important question, let us digress for a moment to consider the following. What is it that allows two entities to 'recognize' one another so that a relationship can form? That is, given a component α and a component β, what is it that allows a relation $* = r/\alpha\beta$? The answer can only be that they mutually specify one another in exactly the same way as an autopoietic form comprising elements which are self-specifying; that is, it is the mutual interaction of the two components. In truth, a relation — a production — is no more than a *description* of a dynamic interplay of self-specifying entities. In self-specification, in the form of a triadic star relationship, the components themselves and the emergent relationship arise. So what, then, about an entity which exists in isolation, i.e., which are not participants in some production (e.g. think of the production $s \longrightarrow \overline{s}\rceil$). It seems that the strong conclusion to which we are led is that the 'recognition' which is apparent in a production is formed by mutual specification. Hence, the components do not have 'semantic' labels attached by which they are recognized and productions are no more than descriptive tools for articulating (or discussing) the dynamic processes of actualization of potential.

So when implementing a system, what can we do? Do we simulate by forming productions based on rules, with a general purpose generic relation and components or is there an alternative? There does appear to be one alternative, but it is wholly impracticable since it requires the creation of a noumenological potential field and, in the light of the arguments in the beginning of the paper, it should be clear we can't do this. We can only simulate. And if we simulate, at what level of granularity do we simulate the productions? Specifically, does it make sense to simulate the $s \longrightarrow \overline{s}\rceil$ production? The answer would appear to be 'no' since there is no 'relationship' characterized by the production; it is just a transformation. Thus, we should opt for the latter implementation of the autopoietic form, with non-trivial (more complex) relations. In the implementation, there will, of course, be many possible actualizations of each production at any instant, depending on what components are available to constitute the relation.

There is one remaining problem which must be addressed in the realization of the simulator of autopoietic systems. Given relations r_i and r_j, if r_i is actualized then we must ensure that it is subsequently only a member of one other relation (despite that all relations are potentially constituents of all other relations). In essence, this is a form of 'bonding' which results from our 'threaded tree' structure. This is done as follows. Allow a relation to communicate an 'actual' tag to all constituents to indicate that it is, in fact, actualized. Thus, if r_i is actualized:

1. if r_i has received no 'actual' tag from any relation, this implies it is potentially a component of *any* relation and consequently output should be to all channels, i.e., to all relations. Otherwise, given r_i received the actual tag from relation j, i.e., it is a component of relation j, output only to relation j.

2. For all components (inputs) which are elements of r_i (i.e. of the production), let these relations know that they are actualized by sending a communication with tag 'actual' to r_j.

What remains, of course, is to implement these productions and to observe what behaviour our instantiation of this simulated autonomous system exhibits.

3 Recapitulation

To conclude, a short recap on the issues we have discussed may be helpful to clarify the position at which we find ourselves upon completion of this essay. We began by identifying one of the key issues in autonomy: that autonomous systems, despite the apparent paradox, are defined in a context and that it is the mutual relevance of the autonomous system and its context, and the mutual specification of the two, by which autonomous systems arise and must be understood. This led us to consider the most fundamental, and meaningful, of contexts: the nature of reality and existence.

After a brief tour through the main philosophical positions on being and existence — ontology — we argued that the ontology one adopts prejudices what one conceives as being possible and actual. In particular, we argued that the pervasive realistic ontologies of, e.g. Locke or Moore, with their 'binary-valued' attributes of existence or non-existence, irrespective of the observer, does not allow for an adequate treatment of autonomous systems. We dubbed this paradigm 'scientific ontology' as it pervades the thinking of so many of modern scientists. We looked too at the more esoteric idealistic ontologies of, e.g. Berkeley or Kant, with their explicit dependence on the observer. We concluded that the distinction between the phenomena and the noumena in idealism offers at least scope for progress in that it does not adopt a pejorative standpoint on what might be the true noumenal nature of reality, as opposed to our perceptions, and experience, of it. We then argued that a relativistic ontology, borrowing greatly from Kant's idealism but also taking on board the validity of realism and the necessity of dealing with phenomenology and personal experience, is what is required for a sound foundation of autonomous systems.

Allowing a relativity in ontology results in a spectrum of being and existence and does away with the 'binary-valued' view-point on existence. It is this spectrum of existence — more or less-real entities — which, in turn, allows for a possibility for entities to have one or other level of existence or being. We then identified *organization* with this scale of existence or, rather, we idenified it as an 'indicator' to a level of existence.

We concluded that since the development, or actualization, of the potential for existence at a certain level, specifically for existence at an autonomous level, concerns the noumenal aspects of entities, and is not at all contingent upon the phenomemology of humans or any other cognitive entity, then this actualization cannot be deterministically invoked by a 'third party' and requires *self-actualization* or *self-organization*. Such self-organizing autonomous systems are effectively life-forms. However, the possibility still exists for the *simulation* of autonomous systems through self-organization in our phenomenological domain, rather than the self-organization of life-forms in a noumenological domain.

We looked at the ramifications for autonomous systems of the effective combination of realism and idealism in this relativistic ontology and the consequent constructivist nature of perception and cognition. This led to the identification of a pathalogical flaw in the development of autonomous systems using conventional representational information-processing approaches: the implicit homunculus.

We looked in detail at the Natural Philosophy of J.G. Bennett, with its relativistic ontology which posits a stratification of existence. We noted that autonomous systems, as we understand them, correspond to levels five through eight, inclusive, in this ontology and thus we are presented with a spectrum of autonomy. It is important to note too that along with this ontology, Bennett presents a methodology (*Systematics*) for dealing with and understanding each of these levels. This methodology hinges upon the correlation between the ontological level and the number of terms which a system of that level must possess in order to exhibit the characteristics of that level. This led us then to consider Varela's and Maturana's concept of autopoiesis as a form of entity at the fifth level of existence — a so-called quinquepotent entity. While *Systematics* provided us with a methodology for determining the boundary conditions of a specific type of autonomous system, it is the calculus of Indications which allows us to contemplate the design of simulations of autonomous systems. We presented some preliminary results of such design, expressed again in terms of indicational forms and we developed a set of necessary conditions for the realization of autopoietic — self-renewing autonomous — systems. Although the simulation system which was described has been validated by realizing a simple allopoietic control system for target tracking, it remains to validate the organizational principles of autopoietic systems and the conditions for the realization of autopoiesis with this simulator. Once this is achieved, we can then proceed to develop simulations of autonomous systems of higher complexity, beginning with self-replicating autonomous entities.

References

[1] Bennett, J.G. *The Dramatic Universe*, Vol. 1, The Foundations of Natural Philosophy, Revised Edition, Claymont Communications, Charles Town, West Virginia, 1956.

[2] Babloyantz, A. *Molecules, Dynamics, and Life*, John Wiley & Sons, New York, 1986.

[3] Bohm, D. *Wholeness and the Implicate Order*, Ark Paperbacks, London, 1980.

[4] Furlong, D. 1986. 'On The Problem of Structure in Perception', Technical Report, Department of Microelectronic and Electrical Engineering, Trinity College, Dublin.

[5] Langton, G. L. (Ed.) *Artificial Life*, Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems, Vol. VI, Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley, New York.

[6] Proceedings of 2nd Conference on Artificial Life, Santa Fe,

[7] Proceedings of the First European Conference on Artificial Life, Paris, 1991, to be published by Bradford Books / MIT Press.

[8] Laudan, L. *Science and Relativism – Some Key Controversies in the Philosophy of Life*, The University of Chicago Press, Chicago, 1990.

[9] Maturana, H. R. and Varela, F. *Autopoiesis and Cognition — The Realization of the Living*, D. Reidel Publishing Company, Dordrecht, Holland, 1980.

[10] Marr, D. 1982. *Vision* W.H. Freeman and Co., San Francisco.

[11] G.E. Moore, 'A proof of an external world' in *Philosophical Papers*, Macmillan, New York, 1962.

[12] Spencer-Brown, G. *The Laws of Form*, George, Allen, and Unwin, London, 1969.

[13] Varela, F.J. *Principles of Biological Autonomy*, Elsevier North Holland, New York, 1979.

[14] Vernon, D. 1991. *Machine Vision*, Prentice-Hall International, London.

[15] Vernon, D. and Furlong, D. 'Limitations of Scientific Ontology', Technical Report, Department of Computer Science, Trinity College, Dublin, Ireland, 1992.

[16] L. Wittgenstein, *Tractatus Logico-Philosophicus* (1921), trans. D.F. Pears and B.F. McGuiness, Routledge and Kegan Paul, London, 1961.