

# Modelling Emergence

---

Francis HEYLIGHEN\*

*PESP, Free University of Brussels, Pleinlaan 2, B-1050 Brussels, Belgium*

*E-mail: Z09302@BBRBFU01.BITNET*

**ABSTRACT.** Emergence is defined as a process which cannot be described by a fixed model, consisting of invariant distinctions. Hence emergence must be described by a metamodel, representing the transition of one model to another one by means of a distinction dynamics. The dynamics of distinctions is based on the processes of variation and selection, resulting in an invariant distinction, which constrains the variety of and thus defines a new system. A classification of emergence processes is proposed, based on the following criteria: amount of variety, internality/ externality of variation and selection, number of levels, and contingency of constraint. It is argued that traditional formal and computational models are incapable of representing the more general types of emergence, but that it is possible to generalize them on the basis of the dynamics of distinctions.

**KEYWORDS:** emergence, models, distinction dynamics, variation-and-selection, variety

## 1. Introduction

The basic difficulty in understanding evolutionary processes in which something qualitatively new is created consists in modelling such novelty within an existing language. The paradox that confronts us is that each language in order to be understandable must consist of conventional, invariant elements (words) and rules for the combination of those elements (syntax, grammar), whereas the phenomenon we wish to describe is by definition more than a combination of existing elements. How can something fundamentally new be expressed in a known framework? Such fundamentally novel phenomena which cannot be reduced to a mere combination of known things are traditionally called "emergent". Hence the problem we are addressing is the problem of how to model emergence.

One way to define emergence is to call a behavior of a system emergent when it can no longer be described by the model that described the system until then (Rosen, 1985). The emergent behavior is novel in the sense that we could not in any way predict it by means of the model we have of the system. A model can in general be conceived as a *set of possible states* that can be observed on the system, together with relational constraints which determine which states are to be expected under which conditions. Usually these constraints have the form of *transition rules*, determining which state(s) can be reached from a given initial state. As long as the evolution can be described by such a sequence of transitions within the given state space there is no need to speak of emergence: this is

---

\* Senior Research Assistant NFWO (Belgian National Fund for Scientific Research)

ordinary mechanical or dynamical evolution. Creative or emergent evolution takes place when *the space of possible states and/or the transition rules change*. The change of the model of a system can be understood as a change of the system itself: the system acquires a new organization and a new behaviour, sometimes even a new identity (or identities, when the number of systems is multiplied).

For example, a billiard ball moving on a billiard table can be perfectly modelled by a mechanistic, Newtonian model. Its variables, determining the possible states, are the 3 coordinates of its position and the 3 coordinates of its momentum, together 6 degrees of freedom (Heylighen, 1990a). Given those variables and the forces, the movement of the ball can be perfectly predicted. However, suppose that the ball would break in two. Now two bodies need to be taken into account, each with 6 degrees of freedom, together  $12 = 6 \times 2$  degrees of freedom. Clearly the space of possible states has changed. Such a change cannot be described within the original, 6-dimensional mechanical model. Equivalently, we might imagine a situation where two balls would stick together, resulting in a decrease of the state space from dimension 12 to dimension 6. More radically, we could imagine that one of the balls would melt, resulting in a liquid with an infinite number of degrees of freedom. All these events are examples of emergence: the system acquires new properties, which cannot be reduced to one of the variables (properties) of the original model. The breaking or melting of balls may appear to be a highly unlikely event, and that explains why Newtonian mechanics is in general an adequate theory for describing rigid objects like billiard balls or planets. Yet, in general, emergence as it was defined here is a very common phenomenon.

Consider for example a chemical substance brought into contact with other chemical substances. At first the system may be described in a state space with a number of variables representing the concentration of each of the present substances. Yet we know that through chemical reactions new substances will be formed, so that new variables need to be introduced. When the chemicals are of a somewhat complicated type, for example proteins and enzymes, the number of newly created substances will be virtually unlimited. Hence the state space will increase to a practically infinite dimension! Another simple example is crystallization: the molecules dissolved in a fluid move independently of each other, and hence their overall state space is the product of the state spaces of all the individual molecules, a virtually infinite number. However, when the solution dries up, the dissolved substance will crystallize, and all molecules will be arranged in a regular, fixed structure, without any freedom to move independently of each other. The number of degrees of freedom has been reduced from infinite to 3 (assuming there is just one crystal positioned in 3-dimensional space). As a last example, all biological systems are "self-modifying systems" (Kampis, 1990) that continuously acquire new components (e.g., cells) and new properties, thus transforming their space of possible states and internal dynamics.

If emergence is defined as a process which cannot be captured by a fixed model, then the question is: is it at all possible to model an emergent process? The solution to the riddle is simple: if emergence is characterized by a change of one model to another one, then we need a model which is able to represent several possible models and their transformations. In other words, a model of emergence must be a *metamodel*: a model, whose states are models, and whose transition rules determine the shift from one model to another one. Hence the theory we need for describing emergence must start from an analysis of what a model is and how it can change.

## 2. Models as distinction systems

What we called properties, variables, or degrees of freedom, namely the attributes which make it possible to describe a variety of possible states, can ultimately be understood as *distinctions* (Heylighen, 1988, 1989a, 1990a), or families of distinctions. Indeed, a property is something which has distinct values: minimally the two values "true" or "false", maximally, in the case of continuous properties such as position, an infinite number of values. Without distinctions there would be no model, since no changes of state could be represented.

However, a distinction also implies an "assimilation": the number of potential distinctions is infinite, the number of actual distinctions or distinction families (e.g. continuous variables) must be finite in order to have a finite model. The essence of a model is indeed that it is simpler than the phenomenon it represents. Otherwise nothing would be gained in creating a model (Heylighen, 1990a). Hence certain potential distinctions will not be incorporated in the model, and phenomena which are distinct in this respect will be assimilated, i.e. put in the same class. For example the Newtonian model of the billiard balls discussed before does not distinguish the color, the composition or the temperature of the balls. No two existing billiard balls are completely the same; yet in the model we assume that they can be described by identical properties and laws of motion.

A distinction is always relative to a "point of view": a particular way of observing or interacting with the system that is to be distinguished. From one point of view two billiard balls may be identical (e.g. they have the same mass and velocity), from another point of view they are distinct (e.g. they have a different color). The first point of view may be that of a color-blind person, or simply that of a Newtonian scientist who is only interested in predicting the movement of the balls, and who hence does not make any explicit observations of visual properties. Distinctions are hence always to an important degree subjective.

Yet they are not arbitrary: the reason the Newtonian researcher is interested in the position and momentum of the ball is because he knows that the values of those properties at a certain instant are sufficient to predict all future values of those properties. No other properties need to be taken into account. The system, consisting of the distinction families *position + momentum*, and the Newtonian laws of motion expressing relations between the values of those distinctions at different moments in time, is *closed*: it is self-sufficient, it does not require any external data in order to completely determine the evolution of these values. Closure of a model is what makes prediction and control of the modelled system possible. Another way of expressing this property is by remarking that all distinctions within the distinction system are *conserved*: two states at a particular instant that are distinct according to the model, will remain distinct for all future times (predictability), and are the result of states that have been distinct at all previous times (reversibility) (see Heylighen, 1989a). In fact closure can in general be defined as the invariance of distinctions (Heylighen, 1990b, 1989b).

So we may conclude that when modelling we will tend to make those distinctions that are as invariant as possible, i.e. that make the model closed. Yet modern science has come to the conclusion that it is not possible to construct closed models that completely capture the behaviour of a system, like in classical, Newtonian mechanics. There are always factors that cannot be predicted or explained, and hence closure or distinction conservation is at best partial. The traditional example of such a non-classical theory is quantum mechanics (Heylighen, 1990a,c). Until now those limitations on closure have

been mostly interpreted in a negative way: as uncertainties, as indeterminacies, as restrictions on our capacity to understand nature.

Yet when we look at the phenomenon of emergence we see lack of closure in the first place as a positive, creative factor, which may help us to understand the complexity of nature as it arises during evolution. Emergence as conceived here is defined not so much as a phenomenon requiring the destruction of a closed model, but as a phenomenon requiring the replacement of one (partially) closed model, by another—possibly more complex—(partially) closed model. Emergence means that new distinctions must be made in order to find a (partially) closed model. Such a partially closed model defines a new system. The distinction between that overall system and its environment (the distinctions outside the closure) summarizes in a sense the emergence process, by representing the novelty that was created by a single concept. Yet that global system may in itself contain a large number of novel subsystems or properties (distinctions of a lower level or order).

We may conclude that a model of emergence must be a model of how new distinctions appear. Such a model may be called a "distinction dynamics". This approach is in sharp contrast with the classical, mechanistic view of science. Indeed, mechanistic models are characterized by the fact that all distinctions are a priori fixed: no change of distinctions is possible (Heylighen, 1990a,c). We will now discuss the basic principles of the dynamics of distinctions, and then apply them in a more detailed classification of different types of emergence.

### 3. Emergence through variation and selection

#### 3.1. Variation

The dynamics of distinctions, as I have started to develop it (Heylighen, 1990a,c), is based on a generalization of the Darwinian principle of variation and selective retention of stable ("fit") systems. Variation means the generation of a variety of simultaneously present, distinct systems (parallel or spatial variety), or of subsequent, distinct states of the same system (sequential or temporal variety). For example, a population of animals with different genomes is characterized by spatial variety. A single animal that attempts to reach a goal by trial-and-error disposes of a temporal variety of possible actions. *Variety* can be defined as a measure of the amount of such distinct states or systems. By "distinct" I mean that they can be distinguished in a relatively stable or invariant way. This means that states or systems that are so unstable or transitory that they cannot be observed or controlled are not considered to be distinguishable or distinct. Variation is hence that part of the process of evolution that makes variety increase, or that produces more distinctions.

The mechanism of variation can be either deterministic, generating new variants in a predictable and systematic manner, or random. As we will see, this does not establish the determinism or indeterminism of the overall process. The mechanism can in general be understood as a (*re*)combination of already existing distinct entities (systems or attributes of systems): either several entities are brought into contact or connected in a certain way, forming an "assembly", or one or more entities are split up into component entities, that may or may not be combined into a new assembly. For example, building a house is a combination of different bricks into an assembly. A chemical reaction requires a splitting up of a one or more types of molecules, and the recombination of parts of one molecule type with part of another molecule type.

We can distinguish two types of recombination here: either the entities to be combined with the original entities already belonged to the system under study (*internal* variation), or the entities come from outside the system under study (*external* variation). For example, in chromosome recombination during sexual reproduction, a second chromosome comes from the sperm of another (male) organism, outside the original genome of the female, and hence can be considered external. However, when the system under consideration is not the individual female but the gene pool of the species, this type of variation must be considered internal.

### 3.2. Selection

The second main concept we need for a dynamics of distinctions is a mechanism that removes distinctions, in other words that decreases variety. This mechanism is selection. It can be defined as the elimination of certain distinct entities (systems or states), reducing the number of remaining entities. The principle that governs "natural" selection can be formulated as a *tautology*: stable entities are retained, whereas unstable ones are eliminated. Although a tautological principle may appear trivial, its value lies in the fact that it can be used as a postulate or axiom in a formal system, which by definition is considered to be always true. Such a formal system (for example classical logic) can be used to derive a lot of very useful and non-trivial consequences ("theorems").

Again selection can be classified as either internal or external: the criterion that decides whether an entity will survive (is stable), may depend on something external (the environment) to the entity, or on the internal structure of the entity. Certain assemblies are intrinsically unstable and hence will fall apart, independently of their environment. On the other hand, the survival of the entity may be contingent upon its context or environment. If it survives, we may say that it is "fit" or "adapted" to its environment.

### 3.3. Interaction of variation and selection

Let us first consider a limit case of variation-and-selection: the case where nothing new is created. This happens when variation and selection perfectly counterbalance each other's action: for each new variant (system or state) generated, the previous variant is eliminated, so that at each instant there is only one possible variant present, and this variant is unambiguously selected out of the variety of possible states. What is generated through variation is selected, and vice versa. This is what happens during deterministic, mechanical evolution. In this situation it is not necessary to conceptually separate variation and selection: the evolution can be modelled by a determined trajectory in a state space (Heylighen, 1990c).

In general, however, variation and selection are *independent*: the variation mechanism "does not know" which of the variants it generates will be selected, and the selection mechanism "does not know" which will be the variant that satisfies its selection criterion. This is obvious in cases where variation is internal and selection external (or vice-versa). For example, in biological evolution the internal variation of the genomes within a species through mutation and sexual recombination is independent of the environmental criteria which determine whether a specific genotype will survive. (This is what distinguishes a Darwinian type of evolution from a Lamarckian one, where the variation of the organism is a function of the environment). However, even in cases where both variation and selection are internal they are in general independent.

An example can be found in classical problem-solving, where a problem is defined by a well-defined goal (selection criterion) and a well-defined search space with

operators for generating new states (variation mechanism). If variation and selection were perfectly correlated, like in classical mechanics, solving the problem would be trivial: it would consist of applying a well-selected sequence of operators (function or algorithm) to the initial state, with the guaranteed result of generating the goal state. In practice, however, things are not that easy. The best one can do is to use heuristic rules which enhance the probability for success, but without guarantee that the variation mechanism will effectively generate a state that satisfies the selection criterion.

Problem-solving could still be considered a case of "artificial" selection where the goal state is determined by an outside user. However, even the natural selection criterion of internal stability does not allow one to predict which system or state to be generated by a completely determined variation process will turn out to be stable. It can be shown that there are perfectly deterministic, computable systems ("cellular automata") whose long term behavior (reaching a stable state or stable limit cycle, or continuing to vary indefinitely) cannot be computed in a finite time: the question "Will the system evolve towards an attractor of a specific type?" is computationally undecidable (Wolfram, 1984). Hence we see that the "creativity" of a process, in the sense of how "unexpected" its result is, does not depend on the degree of determinacy of its underlying variation process, but on the relative independence of variation and selection mechanisms.

### 3.4. Variety and constraint

However, there is another sense in which a process can be creative, namely if it generates a distinction that is more than just an arbitrary assembly or combination of existing distinctions. Like mentioned before, what we recognize as a distinction is something that has a minimum degree of invariance, that is to say that it is conserved and belongs to a system characterized by some form of closure. Now it is possible that an assembly of distinct entities forms a system that in itself is more invariant than its subsystems and their relationships.

For example, our body consists of billions of cells. Continuously there are cells that die, while others are newly produced. Yet our body as a whole maintains a continuous identity, in spite of all these creations and destructions among its components. A living body is said to have an "autopoietic" organization: it rebuilds itself continuously, maintaining an invariant identity (Maturana and Varela, 1980). The distinction between my body and the environment is stable, yet the distinctions between my cells change all the time. What is naturally selected here is not a static combination of cells, but a dynamic organization characterized by an overall distinction: the identity of the organism. The selection that maintains such a global, macroscopic organization operates at a different level than the variation that produces different combinations of subsystems.

What is selected here is not a specific variant (distinction) but a *closed system* of distinctions. Closure can be defined as a property of assemblies of distinctions that determines one (or more) stable, *higher-level* distinctions (Heylighen, 1989b, 1990b). A distinction A can be said to reside at a higher level than another distinction B if it is more invariant, that is to say that B can change or disappear without any change occurring in A or A's relations with outside systems, whereas the destruction of A will necessarily change B and/or B's relations with other systems. For an example, we may look at the human body (A) and one of its cells (B), or at a limit cycle or attractor (A), and one of the states belonging to it (B). Closure can be viewed as a *constraint* on the system's behavior, precluding the variations that perturb the survival or continuity of the system's identity (A), while allowing the variations of B, that have no influence on A's

maintenance. The whole of all allowed variations forms the system's potential variety or freedom.

It is through the concept of closure that we can understand multi-level emergence of new distinctions. The fact that variation and selection generally operate on different levels explains the fact that we usually find them to be independent, even when both are internal to the system. It also explains how an invariant distinction can be retained through selection as a constraint, without stopping the further variation which may go so far as to drastically modify the system without destroying its identity (self-modifying systems, cf. Kampis, 1990). The constraint as well as the remaining variations can be either internal or external to the system.

#### **4. Classification criteria for emergence**

The dynamics of distinctions, being a metamodel, requires a classification of distinctions (varieties) into distinct types (metadistinctions), determining a meta-state space of possible distinction systems (Heylighen, 1988). Let us restate our basic concept of emergence: the transition after variation from a given system to a different, selectively retained system, characterized by a stable distinction, and to be represented by a new, (partially) closed model. This system may or may not contain a number of new subsystems or stable distinctions on one or more lower levels. The discussion of variation and selection has provided us with a number of fundamental attributes or parameters of evolutionary processes. They will allow us to put forth a first taxonomy of emergence types. These attributes are the following: amount of variety, internality/externality of variation and selection, single- or multi-level character of emergence, and type of constraint.

##### *4.1. Amount of variety*

By variety we mean the variety of possible states the new system is capable to attain (potential, temporal variety) or the variety of subsystems the system contains (spatial variety). Both varieties are linked, in the sense that if the subsystems can vary independently of each other (that is to say they are not rigidly connected like in a crystal), then the number of states that the system as a whole can attain is equal to the product of the number of states the subsystems can attain. Hence the larger the number of subsystems, the larger the potential variety of the global system. (the overall temporal variety grows exponentially with the number of subsystems).

We can classify emergence processes according to the absolute amount of variety the new system contains, or according to the increase or decrease of variety with respect to the previous system that has been replaced by the new one. Concerning the first criterion we may roughly distinguish between zero variety, small variety, and large variety. An example of the first type would be a dynamical system that reaches a fixpoint or point attractor: no further evolution is possible any more, except by radically destroying the system. A physical example would be the crystallization of a liquid, where the resulting crystal is embedded in fixed mass (e.g. the crust of the earth). An example with small emergent variety would be a dynamical system that reaches a (one-dimensional) limit cycle: although there is one degree of freedom left (position on the cycle), there is very little variety compared to the original variety in the state space. An example with large variety is boiling or sublimation, resulting in a gas with a practically infinite number of degrees of freedom.

Concerning the second criterion, the change in the amount of variety, we may remark that all examples of emergence or self-organization studied in so-called "complex dynamics", "complex systems" or "dynamical systems" models (e.g. non-linear thermodynamics, cellular automata, chaos theory, ...) correspond to reductions of variety (Heylighen, 1990c). The main concept in these approaches is that of a (strange) attractor. We must remember that, however strange or chaotic, an attractor is still a *part* of an existing state space, which hence, by definition, has a smaller variety than the original state space. The reason that it is impossible to model increases of variety in such approaches, is that one starts from the notion of a "dynamical system" as a predictable, albeit irreversible, process, that is to say a process that can only eliminate distinctions, not create them (Heylighen, 1989a). Physical examples of such processes would again be crystallization, but also the emergence of dissipative structures in thermodynamic systems far from equilibrium.

Examples of variety increase would be again boiling, sublimation or melting, or the examples above of the billiard ball breaking in two, or the chemical reaction in which additional compounds are formed. A biological example would be reproduction, which in the simplest case is just the splitting of a cell in two cells, doubling the number of degrees of freedom.

In general a process will not be characterized by just an increase or decrease of variety: certain types of variety may increase, while others may decrease. For example, a four-legged species of animal that develops wings gains variety, since it can now also fly, that is to say move in a third dimension. However, the transformation of the front legs into wings means that those legs can no longer be used for other actions, e.g. digging a hole, and thus there is a decrease of variety along this other dimension. Similarly, a human being who enters an airplane gains control over a third dimension, but by the same act loses the capacity to go down into caves. However, the invention of the airplane definitely constitutes a gain in the overall variety of actions a human is capable of. Indeed the human still has the choice whether to mount an airplane, or not. This additional variety of choice, however, is situated at another, "meta"-level with respect to the variety of different movements (see further).

#### 4.2. *Internality - Externality*

The variation and selection mechanisms that lead to the emergence of a new system can be classified according to their internal or external origin. The reaching of an attractor for a dynamical system is clearly an example of internal selection after internal variation (both the variation mechanism and the criterion distinguishing stable attractor regions are completely determined by the definition of the system). The adaptation of an organism to its environment by trial-and-error is an example of internal variation (actions of the organism, constrained by its autopoietic organization) with external selection. The growth of a crystal which absorbs more and more molecules from the solution in which it was formed is an example of external variation (the newly added molecules come from outside the crystal) with internal selection (the stable lattice structure which defines the system determines the choice of which molecules are absorbed at which positions). Another example would be a society, e.g. the USA, that selectively allows immigrants from many different countries, thus increasing its (spatial) variety, but without changing the basic organization that defines its stable identity. An example of external variation with external selection would be a child that tries to succeed in school (selection by means of examination) by mimicking or taking over behavioral patterns from its teachers (external variation).



Again we must remark that in general we will find few cases of purely internal or external processes: mostly part of the variation will be internal, part will be external, and equivalently for selection. For example, the general cognitive development of an individual through learning will depend on internal variation (trials, thought), external variation (assimilating ideas from others), internal selection (self-concept, personal criteria for success, goals of the individual) and external selection (rewards and punishments from the environment).

One of the most important functions of the internal-external distinction is to attract attention to the limitations of those models that only consider one type of process: for example, the dynamical systems model which is purely internal, and the model of learning through instruction which is purely external. In fact both these types of models are limited in the same way: they do not recognize the creativity resulting from the interaction between internal and external processes. Like we noted when discussing the relation between variation and selection, it is the *independence* of these subprocesses which engenders the unpredictability or novelty of their interaction products. The unpredictability is especially applicable to external influences, since these are in principle unlimited: they might come from anywhere outside the original system.

It is also important to repeat that internality or externality are not absolute properties, but properties depending on the point of view, i.e. the choice of system boundary (system-environment distinction), which determines what is inside and what is outside. Yet, since we cannot consider the "universe as a whole" as a closed, well-defined system, each system will always have an environment, and hence there will always be external phenomena. In practice, however, it is possible to study "external influences" on emergence within a closed system (e.g. an "artificial world", simulated on a computer, Pattee, 1989), by designing the system such that it consists of (at least) two *independently* evolving parts capable of interaction. One part can then provide external variation and selection for the other part. In computer simulations this would mean that the different parts of the simulation evolve through different variation mechanisms and (internal) selection criteria, which are open for external influences.

#### 4.3. Number of levels

Creative processes can be classified according to whether they involve one level, two levels, or multiple levels. An example of a one-level process would be simple problem-solving, for example an animal that is searching for food. The animal must produce a number of actions (internal variation, for example looking for food at different locations), eliminate those that are not successful, and retain the one that results in the desired goal. The action that is eventually selected is not intrinsically more invariant than the ones that were eliminated: it is just preferred with respect to the goal at that time, and it will stop as soon as the hunger has been satisfied or the food source has been exhausted. Remark that such a process, although it may be called "creative", does not really correspond to an "emergence" as we have defined it, since no new closed model is needed to describe the situation where the animal has satisfied its hunger. At most we would need to introduce a new goal, but the space of possible states remains the same.

A dynamical system that reaches a limit cycle determines a higher-order distinction between the attractor and its environment, that is more invariant than the distinction between the different states within the attractor: the system may still undergo transitions from one attractor state to another attractor state, but will not leave the attractor. On the other hand, if the attractor itself would disappear this would mean at the least that the states that were originally inside the attractor could now undergo transitions to states

outside it. Hence their relations with other states would change fundamentally. The attractor is closed with respect to the states. This is an example of a two-level process.

A multi-level process is characterized by the fact that the newly emerging, closed distinctions from the second level can again undergo recombination, resulting in closed distinctions of a third level, and so on. The closed distinctions at different levels form a partial order, which can locally look like a hierarchy or tree (Heylighen, 1989b). An example of a multi-level structure is the human body: cells are subsystems of organs or circuits, which are subsystems of the body. Notice that the same subsystem (e.g. a cell from the heart) can belong to several higher-order systems (e.g. the sanguine system and the musculatory system).

It must be noted that the traditional "dynamical systems" models, and their discrete counterparts simulated on computer, are incapable of modelling multi-level emergence, since they do not represent the emerging higher-order distinctions (attractors) in the same way as the lower-order distinctions (states) out of which they have emerged. Hence the process of emergence of a closed distinction cannot repeat itself at a higher level: the simulation stops after a single creative step. The reason is that such models are not metamodels: they do not treat closed systems of distinctions ("models") as elements or states of a higher-order model.

#### 4.4. *Contingency of constraint*

In a multi-level system the closed distinction of the highest level can be conceived as a constraint on the further variation of the system. Constraints can be classified according to their strength, but this is equivalent to a classification based on the amount of remaining variety, which was discussed before. It is more interesting to distinguish different mechanisms of constraining the system's variation. On the most abstract plane, it is possible to distinguish types of closures characterized by different mathematical properties (Heylighen, 1990b).

Another important classification criterion is the *contingency* of a constraint: either the constraint is absolute and fixed, precluding and allowing always the same variations, or different variations will be selected in different situations, depending on the circumstances. This second, contingent type may be called "control" rather than "constraint". It is more active, flexible and adaptive than the first type. For example, an attractor or a crystal symmetry are fixed constraints. Knowledge on the other hand is a control on actions that will select different actions in different circumstances. We might also distinguish the two types by calling the former, absolute constraint "structural", since it fixes the organization of the system, and the latter, contingent one "functional", since it operates in function of a situation, without presupposing a fixed structure.

#### 4.5. *Metasystem transitions*

Let us discuss a more specific type of emergence characterized by a contingent constraint: the metasystem transition. The emergence of life, of multicellular organisms, and of our own human intelligence are examples of metasystem transitions. Such a transition was defined by Turchin (1977) as the multiplication and integration of systems with the formation of a higher-order system controlling them (see also Heylighen, 1990d). Let us try to classify this concept in our provisional taxonomy.

This is obviously a two-level process, with a second level embodied by the control system. (we must remark that metasystem transitions can occur subsequently: the control systems of the second level can again be multiplied, integrated and controlled by a

control system of the third level, and so on.) The multiplication and integration of subsystems clearly corresponds to an increase of the variety of the global system. To what corresponds Turchin's concept of a control system in the present terminology? A control system is supposed to be capable of either initiating or prohibiting variations in the systems of the lower level. Hence it can be conceived as an internal selector of potential, internal variations (remark that the initiation of a variation can be seen as the result of preventing all other possible processes, including the process where nothing changes).

Moreover this selection is carried out contingently, in function of the goal that the system tries to reach, and of the specific circumstances which might perturb the reaching or maintenance of the goal state. (More specifically, selection occurs in function of the difference between the (external) situation and the internal goal.) In every system created by natural selection the implicit goal is survival of the system, i.e. invariance of the defining distinction. The perturbations correspond to external variations that might destroy the system. What the system must do is to distinguish the perturbations, anticipate their potentially destructive effects and counteract them before they destroy the system. Hence the system's internal selections will be functions of external variations.

Those variations, if not compensated, would themselves eliminate the system, thus operating an external selection between systems that do adequately compensate the perturbations, and those that do not. The system's control, performing internal selections in function of external variations, can hence be conceived as a *vicarious selector* of internal variations (Campbell, 1974; Heylighen, 1990d), anticipating and substituting for the external selection that might destroy the system. The control can be conceived as a kind of "buffer" or "interface" between internal and external variations and selections, allowing the system to maintain a continuous identity in spite of great internal and external variation. The presence of such vicarious selectors is typical for biological systems, characterized by an elementary cognition (knowledge and perception) of their environment. The "vicariousness" stands for the fact that knowledge "*represents*" external variation and selection, and this is the origin of symbolization and semantics (Heylighen, 1990a). A vicarious selector can itself appear through simple internal variation of potential control systems (possibly internally guided by closure principles) and external selection of the one that adequately substitutes for the environment's selective action (Campbell, 1974).

The relationship between the amount of internal and of external variation can be understood from the law of requisite variety (Heylighen, 1990d). If the internal variety of the system is insufficient to compensate the external perturbations, the system can undergo a new metasystem transition in order to increase its amount of internally controlled variety.

As an example of a metasystem transition we may consider the process of modelling emergence itself. What we called a metamodel is in fact a metasystem whose object systems are models. The increase of variety is produced by considering all possible models instead of just one. The control at the higher level consists in the generation and selection of that particular model which will adequately describe the aspects of the newly created system we are interested in. This control is contingent upon the emergence process that is going on and the goal we have in mind when modelling it.

## 5. Formalizing and operationalizing the model

We have proposed a conceptual framework for analysing and classifying processes of emergence. In order to develop a real scientific model of emergence, it would be preferable to formalize and operationalize the framework, removing remaining ambiguities, and specifying how the model can be used in practical situations. Let us first discuss the shortcomings of existing formal and computational models of emergence.

We have alluded several times to the "dynamical systems" paradigm for modelling emergence, and its discrete implementation in the form of a computer simulation of an artificial world (including cellular automata, artificial life and some models in artificial intelligence). Although in the present definition these models do represent one specific type of emergence (from a state space to an attractor), they are severely restricted, in the sense that they exclude processes characterized by any of the following properties: increase of variety, external variation or selection, and multi-level emergence. My contention is that these deficiencies are due to an insufficiently developed, too mechanistic, conceptual framework, and not to inherent limitations of mathematical formalization or computer simulation. Let me sketch an alternative way to develop formal and computational models.

The model we wish to develop should be "open" in the sense that it is always possible to have new distinctions appear, either externally, by adding new elements, or internally, by the creation of new distinctions through closure. There should not be any limitation on this open-endedness: it must be possible to create any type of new distinction at any moment. Therefore we need a system without a priori restrictions on the type or number of elements. It also means that the elements of the system should not have a fixed meaning, since the appearance of a new distinction may completely change the function or significance of a given element participating in the definition of the distinction. There should not be any "primitive" distinctions, in the sense of distinctions that are used to recursively generate all other distinctions but that are unable to change themselves.

I have proposed the fundamentals of a formal language with those features, based on the concept of a "bootstrapping" definition of distinctions (Heylighen, 1990b, 1991b). The descriptions of the language consist of directed graphs or networks of connected "arrows". The language is open in the sense that more elements (arrows) can be added to a description, changing the distinctions within that description (model). It is also possible to simplify the model by taking out elements. Even when no elements are inserted or removed from the outside, there are closure operations that will internally generate new distinctions by bootstrapping (remaining on the same level: element identification, see Heylighen, 1991a) or by recursion (producing higher levels: element integration, see Heylighen, 1991a).

Two applications of this language have been conceived until now: 1) A reconstruction of the foundations of physics, viewed as the emergence of fundamental geometrical and physical structures (temporal precedence, space-time topology, causality, particles, ...) out of "empty", "meaningless" arrows representing primitive processes (Heylighen, 1991b).

2) The design of a computer support system for structuring complex problem domains, where provisional formulations of the problem are represented as a hypermedia network of chunks connected by different types of links (Heylighen, 1991a). The network is continuously changing, letting new concepts and relations emerge, through the interaction between the human user, who introduces new ideas (external variation), and accepts or rejects proposals from the computer (external selection), and the computer

system itself, which examines different combinations of chunks (internal variations) and selects those that are closed (internal selection), proposing them as potential new distinctions to the user.

Though here the role of the environment (external variation and selection) is played by the user, a similar computer system could be conceived without a user interacting with it. For example, the system might be implemented as an autonomous robot equipped with sensors and effectors that provide feedback from the "real world". Here it should be observed that the sensor must be complex enough to allow a detailed view of the environment, which can be reprocessed by the robot in order to make different distinctions than the ones originally built into the hardware. In other words the robot should have the capacity to conceive new "measurements", "observables" or "semantic functions" (Cariani, 1989; Pattee, 1989), by reorganizing its sensory distinctions on the basis of internal closure and external fit.

It is also possible, like we remarked before, to simulate external variation and selection by having several independently evolving but interacting computational models of emergence (Heylighen, 1989b). Some of these simulated systems could replicate themselves, producing off-spring that is similar but not necessarily identical to their parent system. This would provide a model of increase of variety, since the number of independently evolving systems would increase without a diminishing of their internal variety. By interaction those systems might form closed combinations, determining a higher-order system, which could again start interacting with other systems of the same or different levels, possibly resulting in a system of an even higher level. In that way not only the variety but also the *complexity* of the simulation would increase (Heylighen, 1990b). Finally, the model should be able to develop contingent constraints, vicariously representing external processes. This can happen by simple internal variation with external selection, possibly mediated by internal connections ("bucket brigade" algorithms), as it is being simulated on computer in so-called classifier systems (Holland et al., 1986; Wilson, 1987).

## 6. Conclusion

We have discussed the problem of how to model emergence, where emergence is defined to be a process that cannot be captured by a fixed model. In order to escape from this paradox we must relinquish the mechanistic approach, which requires models to be closed, that is to say to be composed from invariant distinctions. Instead we need an approach that attempts to model how distinctions and hence models change. In other words, we must design a metamodel. A model is based on a space of possible states that are distinguished, and on transition rules between the states. Hence we should start the construction of a metamodel by distinguishing different types of models (metadistinctions, leading to metastates), and by distinguishing different types of transitions between models (metarules). Different models define different systems, and different transitions between models define different types of emergence.

The approach that was followed, the "dynamics of distinctions", is based on a distinction between variation (creation of distinctions or variety) and selection (destruction of distinctions). In general, variation and selection are independent, and hence the result of their interaction cannot be predicted. When they operate on a different level they result in the creation of a new system, characterized by an invariant distinction, defining the identity of the system, which constrains the remaining variation, and thus determines the potential variety of the system.

This scheme leads to the following attributes for classifying emergence processes: amount of variety of the created system, change in the amount of variety during the process, externality or internality of the variation and the selection, number of levels created during emergence or present in the created system, and type of constraint maintaining the system's identity (absolute or contingent). Finally an example of a very fundamental type of emergence was considered: the metasystem transition. It is characterized by increase of variety, appearance of a new level, and contingency of the resulting constraint. The development of a metamodel itself is an example of a metasystem transition with respect to the classical, mechanistic system of modelling.

The further development of this metamodel requires a formalization and an operationalization. The difficulties of traditional formal and computational representations of emergence are due to the limitations of their mechanistic presuppositions. As an alternative, a formal language was sketched in which new distinctions can appear without any restrictions. This formalism can be used to design computer models of emergence, that can be applied in supporting the structuring of complex problem domains by a human user. It was argued that simulations of emergence can also be realized by means of autonomous robots, and by means of independent but interacting computational processes.

A further development of the model will require a better integration between these formal and computational techniques, and the general ideas about variation and selection that were put forth in this paper. This should lead to a general framework in which all the main types of emergence processes can be classified, together with the constraints allowing or precluding specific transitions, simulated, and applied in the design of new models.

## References

- Campbell D.T. 1974. "Evolutionary Epistemology", in: *The Philosophy of Karl Popper*, Schilpp P.A. (ed.). Open Court Publish., La Salle, Ill., p. 413-463.
- Cariani P. A. 1989. *On the Design of Devices with Emergent Semantic Functions*. PhD Dissertation, SUNY-Binghamton, Binghamton NY.
- Heylighen F. 1988. "Formulating the Problem of Problem-Formulation", in: *Cybernetics and Systems '88*, Trappl R. (ed.). Kluwer Academic Publishers, Dordrecht, p. 949-957.
- Heylighen F. 1989a. "Causality as Distinction Conservation: a theory of predictability, reversibility and time order", *Cybernetics and Systems : An International Journal* 20, p. 361-384.
- Heylighen F. 1989b. "Self-Organization, Emergence and the Architecture of Complexity", in: *Proc. European Congress on System Science. AFCET, Paris*, p. 23-32.
- Heylighen F. 1990a. *Representation and Change. A Metarepresentational Framework for the Foundations of Physical and Cognitive Science*. Communication & Cognition, Gent.
- Heylighen F. 1990b. "Relational Closure: a mathematical concept for distinction-making and complexity analysis", in: *Cybernetics and Systems '90*, R. Trappl (ed.). World Science Publishers, Singapore, p. 335-342.
- Heylighen F. 1990c. "Classical and Non-classical Representations in Physics I", *Cybernetics and Systems* 21, p. 423-444.
- Heylighen F. 1990d. "Cognitive Levels of Evolution: from pre-rational to meta-rational", in: *Proc. 8th Int. Conf. on Cybernetics and Systems*, C.N. Manikopoulos (ed.). WOSC, New York.
- Heylighen F. 1991a. "Design of a Hypermedia Interface Translating between Associative and Formal Representations", *International Journal of Man-Machine Studies* (in print).
- Heylighen F. 1991b. "A Structural Language for the Foundations of Physics", *International Journal of General Systems* (in print).
- Holland J.H., Holyoak K.J., Nisbett R.E. & Thagard P.R. 1986. *Induction : processes of inference, learning and discovery*. MIT Press, Massachusetts.
- Kamps G. 1990. *Self-Modifying Systems: a new framework for dynamics, information and complexity*. Pergamon, New York.
- Maturana H. and Varela F. 1980. *Autopoiesis and Cognition : the realization of the living*. Reidel, Dordrecht.
- Pattee H. 1989. "The Measurement Problem in Artificial World Models", *BioSystems* 23, p. 281-290.
- Rosen R. 1985. *Anticipatory Systems*. Pergamon, Oxford .
- Turchin, V. 1977. *The Phenomenon of Science*. Columbia University Press, New York .
- Wilson S.W. 1987. "Classifier Systems and the Animat Problem", *Machine Learning* 2, p. 199-228.
- Wolfram S. 1984. "Universality and Complexity in Cellular Automata", *Physica D* 10, p. 1.