

# The Self-organization of Time and Causality: steps towards understanding the ultimate origin

Francis Heylighen,  
Evolution, Complexity and Cognition Group,  
Vrije Universiteit Brussel  
fheyligh@vub.ac.be

**Abstract:** Possibly the most fundamental scientific problem is the origin of time and causality. The inherent difficulty is that all scientific theories of origins and evolution consider the existence of time and causality as given. We tackle this problem by starting from the concept of self-organization, which is seen as the spontaneous emergence of order out of primordial chaos. Self-organization can be explained by the selective retention of invariant or consistent variations, implying a breaking of the initial symmetry exhibited by randomness. In the case of time, we start from a random graph connecting primitive "events". Selection on the basis of consistency eliminates cyclic parts of the graph, so that transitive closure can transform it into a partial order relation of precedence. Causality is assumed to be carried by causal "agents" which undergo a more traditional variation and selection, giving rise to causal laws that are partly contingent, partly necessary.

**Keywords:** self-organization, cosmology, ontology, time, causality, order.

## 1. The problem of origins

Without doubt, the most difficult and fundamental problem in cosmology is the origin of the universe. One reason why this problem is so difficult is that *all* traditional physical theories assume the existence of time and causal laws. These theories include Newtonian mechanics, quantum mechanics, relativity theory, thermodynamics, and their various combinations, such as relativistic quantum field theories. In all these theories, the evolution of a system is reduced to the (deterministic or more rarely stochastic) change of the system's state  $s(t)$  according to a given causal law (which is typically represented by the Schrödinger equation or some variation of it) [Heylighen, 1990b]. The time  $t$  here is seen as a real number, which therefore by definition takes values between minus infinity and plus infinity. The "system" therefore is assumed to have existed indefinitely. If we apply this same formal representation to the evolution of the universe, then we can only conclude that this universe cannot have an origin at any finite time  $t_0$ , because that would assume that before  $t_0$  there was no system that could evolve, and therefore no previous state that could causally give rise to the "origin" state  $s(t_0)$ . Yet, the observation by Hubble that the universe is expanding, when extrapolated backwards, leads to the conclusion that the universe started at a single point in time, the Big Bang.

The deeper reason for this paradox is that time and causality are part of the ontology—i.e. the set of *a priori* postulated entities—that physical theory uses for representing all phenomena. They therefore cannot be explained from within the

theory. This assumption of the *a priori* existence of time and causality is in fact merely a formalization of our intuition that every moment was preceded by another moment, and that for every effect there is always a cause. In earlier times, this paradox could only be resolved by postulating a supernatural origin: God as the "prime mover" or "uncaused cause" of the universe. This is of course not acceptable in a scientific theory. Moreover, it merely pushes the difficulty a little further, since we still cannot explain the origin of God. Present-day cosmology evades the problem by viewing the origin of the universe as a "singularity", i.e. a point in time where continuity, causality and natural law break down. However, existing theories by their very nature cannot tell us anything about the nature or origin of this singularity, and therefore the explanation remains essentially unsatisfactory.

This problem requires a radical overhaul of existing theoretical frameworks. Recently, a number of alternative approaches have been proposed that may offer the beginning of an answer to the origin of time and causality. These include postulating an imaginary time from which "real" time would emerge [Hawking, 1988; Deltete & Guy, 1996, Butterfield, Isham & Kensington, 1999], process physics, which sees space and time self-organizing out of a random information network [Cahill, 2003, 2005; Cahill, Klinger & Kitto, 2000], the emergence of causal sets from a quantum self-referential automaton [Eakins & Jaroszkiewicz, 2003], and a structural language for describing the emergence of space-time structure [Heylighen, 1990a]. These proposals are heterogeneous, based on advanced, highly abstract mathematics, and difficult to grasp intuitively. They moreover all start from highly questionable assumptions. As such, they have as yet not made any significant impact on current thinking about the origin of the universe.

The present paper attempts to approach the problem in a more intuitive, philosophical manner, instead of immediately jumping to mathematical formalism, as is common in physical theory. To achieve that, we will look at the emergence of time and causality as a process of self-organization, albeit a very unusual one in that it initially takes place outside of time.

## 2. Generalized self-organization

Models of evolution and complex systems have taught us quite a bit about the phenomenon of self-organization, which can be defined most simply as the spontaneous appearance of *order out of chaos* [Prigogine & Stengers, 1984; Heylighen, 2001]. Extended to the level of the universe, this harkens back to the old Greek idea that *cosmos* emerged from *chaos*, an idea that predates more recent metaphysical theories where the cosmos is created by the pre-existing order or intelligence embodied in God.

Chaos here refers to randomness or disorder, i.e. the absence of any form of constraint, dependency or structure. Since maximum disorder is featureless and therefore indistinguishable from emptiness or vacuum, the existence of disorder does not need to be explained. In fact, modern physical theories conceive the vacuum precisely as a turbulent, boiling chaos of quantum fluctuations, continuously producing virtual particles that are so short-lived that they cannot be directly observed. Moreover, physical theory in principle allows the emergence of stable matter out of these quantum fluctuations without contradicting the law of energy conservation:

in quantum theory, particles can be created out of energy in the form of particle/antiparticle pairs. But that just raises the question of where the energy came from. The answer is that the total energy of the universe is exactly zero. The matter in the universe is made out of positive energy. However, the matter is all attracting itself by gravity. Two pieces of matter that are close to each other have less energy than the same two pieces a long way apart, because you have to expend energy to separate them against the gravitational force that is pulling them together. Thus, in a sense, the gravitational field has negative energy. In the case of a universe that is approximately uniform in space, one can show that this negative gravitational energy exactly cancels the positive energy represented by the matter. So the total energy of the universe is zero. (Hawking, 1988, p. 129)

What we need to explain further is how such separation of positive and negative energy can occur, i.e. how the initially homogeneous chaos can differentiate into distinct spatial regions, particles and fields. Numerous observations of chemical, physical, biological and sociological processes have shown that some form of order or organization can indeed spontaneously evolve from disorder, breaking the initial homogeneity or symmetry. The only ingredients needed for the evolution of order are random *variation*, which produces a variety of configurations of the different elements, and the *selection* of those configurations that possess some form of intrinsic stability or invariance. The selection is natural or spontaneous in the sense that unstable configurations by definition do not last: they are eliminated by further variation. The stable ones, on the other hand, by definition persist: they are selectively retained. In general, there exist several stable configurations or "attractors" of the dynamics. However, random variation makes that the configurations will eventually end up in a single attractor, excluding the others.

This is the origin of *symmetry breaking*: initially, all attractor states were equally possible or probable (homogeneity or symmetry of possible outcomes); eventually, one has been chosen above all others (breaking of the symmetry). What forces the symmetry breaking is the instability of the disordered configuration: this initially homogeneous situation cannot last, and a "decision" needs to be made about which stable configuration to replace it with. A simple example is a pencil standing vertically on its tip. This position is very unstable, and the slightest random perturbation, such as few air molecules more bumping into it from the left rather than the from right, will push the pencil out of balance so that it starts to fall, in this case towards the right. It will end up lying flat on the right-hand side, thus breaking the initial symmetry where it was poised in an exact balance between left and right. More generally, an initial random fluctuation will normally be amplified by positive feedback until it pulls the whole system into a particular attractor [Heylighen, 2001].

Perhaps counter-intuitively, more variation or disorder produces faster self-organization and therefore more order. This is the principle of "order from noise" [von Foerster, 1960], or "order through fluctuations" [Prigogine & Stengers, 1984; Nicolis & Prigogine, 1977]. The explanation is simple: more variation means that more different configurations are explored in a given lapse of time, and therefore the probability to end up in a stable configuration in that period of time becomes greater. We may conclude that the emergence of differentiated order from initially homogeneous disorder is a simple and natural process that requires no further justification. It implies that we can explain the emergence of order out of chaos without need to postulate a pre-existing order or designer.

However, this variation-and-selection mechanism cannot as yet be used to explain the emergence of time, since it assumes processes taking place in time. To

tackle this problem, we need to "abstract away" the notion of time from the two basic components of the process of self-organization, thus arriving at the following generalized notions:

- *generalized variation* does not require change of a configuration in time, but can be a static feature. The only thing needed is the presence of a variety or diversity of configurations. These can be generated by random variations on a simple "template".
- *generalized selection* does not require selective retention, where some configurations are "killed off" or eliminated, while others are allowed to "survive". We only need a selection criterion that allows certain configurations, while making others a priori impossible.

The most general selection criterion that I want to propose here is *consistency*. According to the American Oxford Dictionary, "consistent" has three, related meanings:

- "unchanging in achievement or effect over a period of time": this can be seen as a paraphrase of "stable" or "invariant"
- "compatible or in agreement with something": this can be interpreted as "fitting" or "adapted". For a system to be stable, it needs to fit in or "agree" with its environment, i.e. it should avoid potentially destructive conflict.
- "not containing any logical contradictions": this is the time-independent meaning that we will focus on here; it can be derived from the second meaning by noting that the components of a consistent configuration should be in mutual "agreement".

Consistency in the timeless sense can be seen as a requirement imposed by the law of contradiction in logic: A and *not* A cannot both be true. This law is tautological, which means that it is true by definition. Therefore, it does not need to be justified by recourse to some deeper law or to some external authority, such as God imposing laws on nature. The application of this law in generalized self-organization is that it can be used as a criterion to eliminate configurations that somewhere contain an inconsistency, i.e. some part or aspect of the configuration is in contradiction with some other part of aspect. While this requirement may seem obvious when discussing logical statements, the connection to physical states and self-organization is more subtle.

One illustration I can think of is de Broglie's historical conception of the quantized orbits of electrons around a nucleus. In the spirit of wave mechanics (the precursor of quantum mechanics), these energy eigenstates were seen as closed waves, since the wave has to travel around the nucleus and then connect with it itself. The explanation for quantization was that orbits that are not eigenstates cannot exist because of destructive interference of the electron's wave function with itself. The selection criterion (being an energy eigenstate, or in wave mechanics: being a standing wave with an integer number of nodes) selects specific quantized orbits and eliminates the rest. However, this is not conceived as a process in time, since the non-quantized orbits do not get eliminated one-by-one: they are intrinsically inconsistent, and therefore "logically" unrealizable.

Another example is perception, where the visual system selects a coherent "Gestalt" out of all the possible interpretations of the initial noisy data it receives, and ignores the interpretations that appear inconsistent [Stadler & Kruse, 1990]. Again, the alternative interpretations are not eliminated in a temporal sequence; they simply fail to make sense. Both orbit quantization and Gestalt perception exhibit symmetry

breaking: from the homogeneous mass of potential states or interpretations, they select one (or a few), leaving out the rest.

### 3. The origin of time

Time is in the first place an order relation between events, allowing you to specify whether an event A came either *before* or *after* an event B. Relativity theory has generalized this intuitive notion of a complete or linear order of time by noting that sometimes the order of events cannot be determined: when A occurs outside of the light cone passing through B (which means that it is impossible to send a signal from B that arrives in A or vice-versa), then the temporal order between A and B is indeterminate. For some observers, A will appear to be in the future of B, for others in the past, or in the present. In general, we may say that A and B cannot be ordered absolutely. Therefore, according to relativity theory the order of time is only partial.

A *partial order* is actually a very simple and common mathematical structure. In fact, any arbitrary relation can be formally converted to a partial order by making the relation transitive [Heylighen, 1990a]. To show how this is done, let us represent this arbitrary relation by the symbol  $\rightarrow$ , which can be taken to mean "connects to", according to some as yet unspecified connection criterion. Adding such a relation to a set of individual nodes  $\{A, B, C, \dots\}$  turns this set into a network. The nodes can be interpreted as some as yet unspecified, primitive "events". We will now perform a *transitive closure* of this relationship or network. This means that if the links  $A \rightarrow B$ , and  $B \rightarrow C$  both exist, then the link  $A \rightarrow C$  is added to the network if it did not exist yet. If it turns out that  $C \rightarrow D$  also exists, then transitive closure means that in a second stage  $A \rightarrow D$  is added as well. This adding of "shortcuts" or "bridges" that directly connect nodes that were indirectly connected is continued until the network has become transitive, i. e. until for every  $X \rightarrow Y$  and  $Y \rightarrow Z$ , there exist a  $X \rightarrow Z$  link. This is a purely formal operation of generalizing the definition of the relation so that it includes indirect links as well as direct ones. It is similar to the extension from the relation between people "is parent of" to the transitively closed relation "is ancestor of", or from the relation between natural numbers "is successor of" to "is larger than".

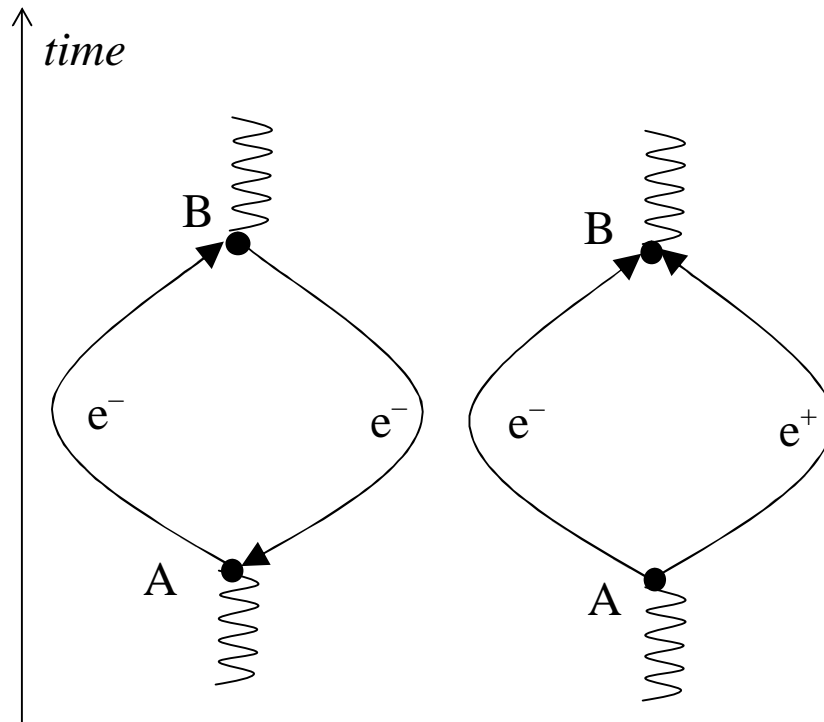
In any relation or network, there are two types of links: *symmetric* (meaning that the link  $A \rightarrow B$  is accompanied by its inverse  $B \rightarrow A$ ), and *antisymmetric* (meaning that the link has no inverse). The combination of transitivity and antisymmetry defines a *partial order* relationship: if you consider only the links without inverse, they impose a clear order on the nodes they link, from "smaller" to "larger", or from "earlier" to "later". The combination of transitivity and symmetry, on the other hand, determines an *equivalence relationship*: if  $A \rightarrow B$ , and  $B \rightarrow A$ , then A and B can be considered "equivalent" with respect to the ordering. If the ordering is interpreted as time, A and B are simultaneous. So, it appears as if this simple transitive closure operation has transformed our arbitrary, random network into a partial order that can be interpreted as an order of time. In other words, we get order (time) out of chaos (a random network).

This is pretty straightforward. However, complications arise if the original relation  $\rightarrow$  (before the transitive closure operation) contains *cycles*. Imagine a long sequence of links:  $A \rightarrow B$ ,  $B \rightarrow C$ ,  $C \rightarrow D$ , ...  $Y \rightarrow Z$ . Transitive closure means that you add all the shortcuts:  $A \rightarrow C$ ,  $B \rightarrow D$ ,  $C \rightarrow E$ , etc. But now you also need to add shortcuts between the shortcuts: if both  $A \rightarrow C$ , and  $C \rightarrow D$  are in the network,  $A \rightarrow D$  also must be added, and so does  $A \rightarrow E$ ,  $A \rightarrow F$ , etc. Eventually, the whole

sequence will be "cut short" by the single link  $A \rightarrow Z$ . This fits in with our intuition about time: if A precedes B, B precedes C, ... and Y precedes Z, then A also precedes Z. But since we started from the assumption that the network is random, the probability is real that it would also contain the link  $Z \rightarrow A$ . In that case, we have found a *cycle*: the sequence of links starting from A returns to its origin. Applying again the transitivity rule,  $A \rightarrow Z$  and  $Z \rightarrow A$  together imply  $A \rightarrow A$ . In other words, A precedes A! This is not grave if we interpret the connection relation  $\rightarrow$  as "precedes or is simultaneous with". The links  $A \rightarrow Z$ , and  $Z \rightarrow A$  are symmetric, and thus they belong to the equivalence part of the relationship. The normal interpretation is therefore one of simultaneity. However, the transitive closure operation implies that *all* elements of the sequence A, B, C, D, ..., Z now become equivalent or simultaneous. This is still not necessarily a problem, since it is principle possible to have many simultaneous events.

The existence of cycles becomes a problem, though, if we make the assumptions that the initial network is both random—because we want order to emerge from chaos—and infinite, or at least unrestricted—because we want the emerging order to represent the infinite extension of time. If we continue to add random nodes and links to the network, sooner or later a very long sequence of ordered nodes will, by the addition of a single link going back to an earlier element of the sequence, turn into a cycle. This cycle, because of the formal operation of transitive closure that is needed to produce an order relation, will turn into an equivalence class. This means that the elements of the sequence, however extended, suddenly all lose their temporal order, and become simultaneous. Simulations of the growth of random networks [Kaufmann, 1995] clearly show that the addition of links will sooner or later connect all nodes into a single cluster or equivalence class. In other words, if we allow the network to grow freely, we will quickly lose our partial ordering and therefore any notion of time.

The only solution seems to be to get rid of the cycles somehow, i.e. to formulate a selection criterion functioning outside of time that excludes cycles and retains only the non-cyclical parts of the random network to constitute the backbone of time. The criterion of consistency is obviously relevant here because cycles in time can lead to the well-known paradoxes of the time machine: what happens if I go back in time before I was born and kill my own father? I have argued earlier [Heylighen, 1990b] that temporal cycles connecting events are either logically inconsistent (A leads to *not* A) or trivial (A leads to A). The trivial cycles merely reaffirm what is already there. The inconsistent ones, on the other, imply that A negates itself and therefore must be null or void. Therefore, selection for consistency would automatically eliminate all such cycles. The effect is similar to the destructive interference undergone by cyclical waves that do not have an integer number of periods: when the wave comes back to its origin with an amplitude opposite to the one it started out with, it effectively erases itself.



The trivial cycles, which do not "self-destroy", on the other hand, are redundant, and can therefore be safely ignored or reinterpreted as partial orders. One way to do this, as suggested in Heylighen [1990a], is to apply Feynman's [1949] interpretation of antiparticles as normal particles moving backwards in time; in other words, we can in principle reinterpret the "back in time" section of consistent cycles as antiparticles moving forward in time, thus changing the orientation of the connections on that section. For example, Fig. 1 shows a simple cycle  $A \rightarrow B, B \rightarrow A$  in which an electron ( $e^-$ ) moves back in time (left). This is reinterpreted (right) as a photon giving rise to a particle-antiparticle pair ( $e^+$ : positron = anti-electron) after which particle and antiparticle mutually annihilate, producing again a photon. Note that in analogy with the quantization of closed waves, such a consistent cycle could be seen as a kind of eigenstate of a hypothetical time reversal operator.

This elimination of cycles leaves us with the non-cyclic parts of the initially random network of connections between events, and therefore with a partial order defining time. Moreover, it can be shown that the remaining connections can be divided in two categories, which can be interpreted respectively as "light-like" (i.e. representing processes with the speed of light), and "particle-like" (i.e. representing processes with a speed lower than light) (Heylighen, 1990a,b). The resulting mathematical structure is equivalent to the *causal structure* of relativistic space-time [Kronheimer & Penrose, 1967], which determines the bulk of space-time geometry. This construction thus not only produces the order of time, but even the fundamental properties of space in its relativistic interpretation (Heylighen, 1990a,b). The argument needs to be fleshed out in much more detail, but already suggests a simple and promising route to a theory of the self-organization of time.

The only additional ingredient we need to recover the full mathematical structure of relativistic space-time is an observer-independent notion of duration, i.e. a unit of time that allows us to measure how much time has passed (Heylighen, 1990b). Given such a unit of time, we immediately get a unit of space or distance for free, since we can define this spatial unit as the distance covered in a unit of time by a signal moving with the (invariant) speed of light.

The existence of invariant time units is equivalent to the assumption that it is possible under certain circumstances for synchronized clocks that are separated and then brought together again to still be synchronized [Sjödín & Heylighen, 1985], because all along they have counted with the same time units. In other words, equal causes (clocks initially showing the same time) produce equal effects (clocks having advanced independently still show the same time). This is actually a problem of causality, which will be discussed in the next section.

#### 4. The origin of causal laws

In relativity theory, causality is usually understood to mean that a cause must necessarily precede its effect. However, this relation of precedence is already fully covered by our notion of time as a partial order between events, and therefore needs no additional explanation.

What remains to be explained is causality in the more traditional sense of "equal causes produce equal effects". This is the sense of causality as a rule or law that allows us to predict which kind of effect will follow given the characteristics of the cause. In my analysis of causality [Heylighen, 1989], I have argued that if we interpret "equal" as "identical" then the principle of causality is tautological, and therefore needs no further explanation. This interpretation corresponds to what I have called "microscopic causality". In practice, however, i.e. in the world of macroscopic observations, when we make predictions we do not assume *identical* causes, but *similar* causes leading to *similar* effects. This interpretation is the principle of "macroscopic causality". The sensitive dependence on initial conditions in non-linear dynamics (the "butterfly effect") and the Heisenberg uncertainty principle, however, both show how similar (macroscopically indistinguishable) causes can lead to dissimilar (macroscopically distinct) effects [Prigogine & Stengers, 1984; Gershenson & Heylighen, 2004]. Therefore, macroscopic causality is not a logical necessity: sometimes the assumption is valid, sometimes it is not. The question that remains then is: why do similar causes *often* lead to similar effects?

A possible approach is to consider a cause-effect relation as a *condition-action rule*,  $A \rightarrow B$ , describing the transition from A (cause) to B (effect): whenever a condition A, i.e. a state belonging to particular subset or category A of world states, is encountered, some agent acts to change this state into a new state, belonging to category B. This perspective fits in with an *ontology of actions* [Turchin, 1993], which sees all change as resulting from a combination of elementary actions performed by one or more agents. An agent in this perspective could be a particle, a field, a molecule, or some more complex system, such as an organism. This implies that causal rules are not absolute or universal, but dependent on the presence of a particular type of causal agent. The presence of this agent functions as a "background condition" necessary for the causation to take place [Heylighen, 1999]. For example, the rule "if a massive object is dropped (cause or condition), it will fall (effect or action)" implicitly requires the presence of gravitation, and therefore the proximity of a mass, such as a planet, big enough to produce gravitational forces. The planet's gravitation here plays the role of the causal agent. In its absence, e.g. in interstellar space, the causal law does not hold.

Such agents—and therefore the laws they embody—are normally the product of evolution. This idea may be illustrated by considering the origin of biological laws.

Living organisms all use the same genetic code, which is implemented by the mechanism of RNA transcription: a particular DNA/RNA triplet is transformed via a



number of intermediate stages into a particular amino acid by the ribosomes and transfer-RNA molecules present in the cell. The causal rules governing this “translation” mechanism together form the *genetic code*. This genetic code is universal, i.e. the same triplet is always transformed into the same amino acid: equal causes produce equal effects. This universality can be explained by the fact that living organisms on Earth have a common ancestor. From this ancestor, all living cells have inherited the specific organization of the ribosomes that perform the conversion from triplet to amino acid. These complexes of RNA and protein were created very long ago by an evolutionary process of self-organization that took place among the autocatalytic cycles of chemical reactions that produced the first living cells. Natural selection has eliminated all variant forms of ribosomes that might have enacted different codes of translation, and thus fixed the present code. Thus, we can explain the law-like character of the DNA code by the selective retention and reproduction of a particular type of ribosomal agents.

Can we generalize such a process of self-organization to explain causal laws in general? The fundamental problem is to explain why natural laws appear to be the same in all regions of the universe. The genetic code example suggests that this may be because all the causal “agents” (which at the lowest level might correspond to elementary particles and fields) had a common origin during the Big Bang, i.e. they are all descendants of the same “ancestors”. However, those original ancestors are likely to have come about contingently, and therefore different universes may well have different laws of nature—e.g. distinguished by the values of their fundamental constants. Why our universe has these particular laws may then be explained by a natural selection of universes picking out the “fittest” or most “viable” universes [Smolin, 1997].

However, the ribosome example suggests that there may have been many alternative laws, enacted by different collections of particle-like agents, that would have been just as effective in generating a complex universe that later gave rise to intelligent life. Biologists have no particular reasons to assume that the present genetic code is the only possible one. While there are arguments based on chemistry to show that the present code is more efficient than most other conceivable codes [Freeland & Hurst, 1998], there is still plenty of freedom in choosing between a large number of codes that appear equally efficient. Biologists assume that these other codes have lost the competition with the present code not because they were intrinsically less fit, but because of contingent events, such as one code being a little more common in the very beginning, which allowed it to profit more from exponential growth to outcompete its rival codes. Here we find again the basic mechanism of symmetry breaking: random, microscopic differences in the initial state (a few more cells with the present code) are amplified by positive feedback until they grow into irreversible, macroscopic differences in the final result. The implication is that there may be a large number of “viable” universes, which all have different laws, but that not all laws are equally viable.

For example, in cosmology the question is regularly raised why in our universe there is such a preponderance of matter over antimatter. The laws of physics as we know them do not exhibit any preference for the one type of matter over the other one. Therefore, we may assume that during the Big Bang particles of matter and of antimatter were produced in practically equal amounts. On the other hand, matter and antimatter particles annihilate each other whenever they interact. This means that such a homogeneous distribution of particles between matter and antimatter states was unstable, and could not continue. It has been suggested that an initial imbalance

between matter and antimatter, which may have been random and tiny, has been magnified by this violent competition between the two states, resulting in the final symmetry breaking, where practically all antimatter was eliminated.

Since antimatter particles are still being formed in certain reactions, we know about the possibility of their existence. However, it is conceivable that the Big Bang witnessed the creation of huge varieties of other, more "exotic" particles, which not only have disappeared since, but which are so alien to the remaining particles that we cannot even recreate them in our particle colliders. Therefore, they are absent in our theoretical models, even as potential outcomes of reactions. Such particles might have embodied very different causal laws, exhibiting different parameters such as mass and charge, and undergoing different types of forces and interactions.

Another implication of this hypothesis is that causal laws may not be as absolute and eternal as physics assumes. If a causal law is "embodied" in a particular type of agent that has survived natural selection, we may assume it to be relatively stable. Otherwise, the agent, and with it the law, would already have disappeared. On the other hand, evolution tells us that no agent is absolutely stable: it is always possible that the environment changes to such a degree that the original agent no longer "fits". This will lead to increased variation and eventually the appearance of new agents that are better adapted to the new environment, thus outcompeting the old ones. When we think about basic physical laws, like those governing the interactions between common elementary particles, such as protons and electrons, it seems difficult to imagine environments where those particles and the laws they embody would no longer be stable. But that may simply be a shortcoming of our imagination, which has no experience whatsoever with totally different physical situations, such as those that might arise inside a black hole or during the Big Bang.

When discussing the contingency of laws it is important to note that there are two types of laws:

- 1) *logically necessary laws*: these are true tautologically, by definition, such as  $1 + 1 = 2$  or the law of contradiction in logic
- 2) *contingent laws*: these could conceivably be different, such as the values of the different fundamental constants in physics.

The difference between these two is not always apparent. Some seemingly contingent laws may in a later stage be reduced to tautologies, which have to be true because of the way the properties that they relate are defined. The law of energy conservation is an example of this: at the most fundamental level, energy appears to be defined in such way that it must be conserved. More precisely, the law of energy conservation, like all other conservation laws, can be derived mathematically (through Noether's theorem) from an assumption of symmetry [Hanca et al. 2004], in this case the homogeneity of time. This means simply that physical processes are independent of the particular moment in time in which they occur: postponing the process to a later moment without changing anything else about the situation will not change the dynamics that takes place. This assumption of time invariance appears to be true by definition: the time coordinate of an event is merely a convention, depending on how we have calibrated our clocks, and should therefore not affect the process itself. In fact, this could be interpreted as another example of the consistency requirement: to be consistent our description of a dynamical process should not change if we merely shift the time coordinate over an arbitrary amount, since time is defined relatively as a precedence relation, and not absolutely, as a number.

At present, most physicists seem to assume that the values of the fundamental constants are contingent, and therefore need to be explained by a combination of

random variation and a selection mechanism such as the Anthropic principle [Carr & Rees, 1979; Barrow & Tipler, 1988] or cosmological natural selection [Smolin, 1997]. However, we must remain open to the possibility that they are necessary, and derivable from some as yet not clearly formulated first principles [see e.g. Bastin et al. 1979, Bastin & Kilmister, 1995 for an attempt at deriving fundamental constants from combinatorial principles]. The example of the origin of the genetic code may remind us that some aspects of a law may be purely the result of chance, while others represent intrinsic constraints that determine which variants will be selected. That selection itself may happen in time, e.g. during a sequence of universes reproducing themselves as envisaged by Smolin (1997), or outside time, by a requirement of consistency like the one we discussed before or like the one that is implicit in symmetry-based derivations of laws based on Noether's theorem.

## 5. Conclusion

The problems of the origin of time and of causality are perhaps the most fundamental of all scientific problems, since all other scientific concepts and theories presuppose and therefore depend on the existence of time and causality. It therefore should not surprise us that as yet no convincing approaches to these problems have been proposed. However, rather than taking time and causality for granted, as practically all theories have done until now, the present paper has argued for a further investigation of these problems.

I have suggested to start from the by now well-documented notion of self-organization, because this concept proposes a concrete mechanism for the emergence of order out of chaos. When considering the origin of the universe, chaos should here be understood in its original, Greek sense, as a total disorder that is so much lacking in structure that it is equivalent to nothingness. Time and causality, on the other hand, are characterized by order. For time, this means the partial order relation of precedence that connects different events while establishing an invariant distinction between past and future. For causality, the order is in the invariance of cause-effect relationships, as expressed by the "equal causes have equal effects" maxim. Invariance can be conceived as stability under certain transformations. Stability can be explained as the result of a process of variation followed by selection that spontaneously eliminates unstable variations. Since chaos automatically implies variation, we only need to explain selection: why are only some of the variations retained?

In the case of causality, the variations can be conceived as causal agents that embody different condition-action or cause-effect rules. In the case of basic laws of physics, the agents are likely to represent elementary particles or fields. Since the agents interact, in the sense that the effect of the one's action forms an initial condition or cause for another one's subsequent action, they together form a complex dynamical system. These systems are known to necessarily self-organize [Ashby, 1962; Heylighen, 2001], in the sense that the overall dynamics settles into an attractor. This means that certain patterns of actions and agents are amplified by positive feedback until they come to dominate, suppressing and eventually eliminating the others, and thus breaking the initial homogeneity or symmetry in which all variations are equally probable. As yet, we know too little about the dynamics of such a primordial complex dynamical system to say anything more about what kind of causal rules might emerge from such a self-organization at the cosmic scale. However, the general notion of self-organization based on variation and

selection suggests some general features of the resulting order, such as the fact that it will be partly contingent, partly predictable, and context-dependent rather than absolute.

In the case of time, this notion of self-organization needs to be extended in order to allow variation and selection to take place outside of time. For variation, this poses no particular problem, since selection can operate equally well on a static variety of possibilities. For selection, we need to replace the dynamic notion of stability as a selection criterion by the static notion of consistency. Consistency can be understood most simply as an application of Aristotle's law of contradiction—which states that a proposition and its negation cannot both be actual. In the case of time, consistency allows us to have a partial order of precedence emerge out of a random graph by eliminating cycles. The connections forming the random graph or network can be interpreted as elementary actions or processes that lead from one event to another. These random links and their corresponding nodes (events) form the initial chaos or variation out of which the order of time is to emerge.

The formal operation of transitive closure transforms a random network into a relation that is partly a partial order, partly an equivalence relation. The equivalence relation encompasses all the parts of the graph that are included in cycles. However, in an infinite random graph, this means in essence the whole graph, implying that no partially ordered parts are left. Therefore, we need a selection criterion that eliminates cycles. This can be motivated by generalizing the paradox of the time machine: temporal cycles that produce actual changes are a priori inconsistent, and therefore "self-negating", like the cyclic waves that undergo destructive interference with themselves. Therefore, we can exclude them a priori.

In both cases—the self-organization of time and of causality—the present description is still very sketchy, applying general principles at a high level of abstraction, but remaining awfully vague as to what the "agents", "connections" or "events" precisely are, or what properties they are supposed to have. At this stage of the investigation, such vagueness is probably unavoidable. However, by proposing a relatively simple and coherent explanation based on the well-understood concept of self-organization, the present approach at least provides some steps towards understanding these fundamental questions. I hope that other researchers may pick up these threads and weave them into a graceful fabric of understanding.

## 6. References

- Ashby W. R. 1962 Principles of the self-organizing system. In H. Von Foerster and Jr. G. W. Zopf, editors, *Principles of Self-Organization* Pergamon, p. 255–278.
- Barrow, J. D., and F. J. Tipler. 1988. *The Anthropic Cosmological Principle*. Oxford University Press.
- Bastin, T., and C. W. Kilmister. 1995. *Combinatorial physics*. World Scientific, River Edge, NJ.
- Bastin, T., H. P. Noyes, J. Amson, and C. W. Kilmister. 1979. On the physical interpretation and the mathematical structure of the combinatorial hierarchy. *International Journal of Theoretical Physics* 18, no. 7: 445–488.
- Butterfield, J., Isham, C.J. & Kensington, S., 1999. On the Emergence of Time in Quantum Gravity. Arxiv preprint gr-qc/9901024.
- Cahill, R.T., 2003. *Process Physics*. Process Studies Supplement, 1-131.
- Cahill, R.T., 2005: *Process physics: from information theory to quantum space and matter*, Nova Science Pub., NY.

- Cahill, R.T., Klinger, C.M. & Kitto, K., 2000. Process Physics: Modelling Reality as Self-Organising Information. Arxiv preprint gr-qc/0009023.
- Carr, B. J., and M. J. Rees. 1979. The anthropic principle and the structure of the physical world. *Nature* 278, no. 605: 230.
- Deltete, R.J. & Guy, R.A., 1996. Emerging from imaginary time. *Synthese*, 108(2), 185-203.
- Eakins, J. & Jaroszkiewicz, G., 2003. The origin of causal set structure in the quantum universe. <http://arxiv.org/abs/gr-qc/0301117>
- Feynman R. P. (1949), The Theory of Positrons, *Physical Review* 76, 749 - 759
- Freeland SJ, Hurst LD 1998. "The genetic code is one in a million". *Journal of Molecular Evolution* 47 (3): 238–48.
- Gershenson C. & F. Heylighen 2004. How can we think the complex? in: Richardson, Kurt (ed.) *Managing the Complex Vol. 1: Philosophy, Theory and Application*. (Institute for the Study of Coherence and Emergence/Information Age Publishing)
- Hanca J., Tulejab S. & Hancovac M. 2004: Symmetries and conservation laws: Consequences of Noether's theorem, *American Journal of Physics*, Vol. 72, No. 4, pp. 428–435
- Hawking, Steven. 1988. *A Brief History of Time*. Toronto: Bantam.
- Heylighen F. 1990a: "A Structural Language for the Foundations of Physics", *International Journal of General Systems* 18, p. 93-112
- Heylighen F. 2001: "The Science of Self-organization and Adaptivity", in: L. D. Kiel, (ed.) *Knowledge Management, Organizational Intelligence and Learning, and Complexity*, in: *The Encyclopedia of Life Support Systems ((EOLSS))*, (Eolss Publishers, Oxford). [<http://www.eolss.net>]
- Heylighen, F. 1999. Advantages and limitations of formal expression. *Foundations of Science* 4, no. 1: 25-56.
- Heylighen, F., 1989. Causality as distinction conservation. A theory of predictability, reversibility, and time order. *Cybernetics and Systems*, 20(5), 361-384.
- Heylighen, F., 1990b. *Representation and Change*. (Communication & Cognition, Ghent).
- Kauffman, S. A. 1995. *At Home in the Universe: The Search for Laws of Self-organization and Complexity*. Oxford University Press, USA.
- Kronheimer E. H. & R. Penrose 1967, "On the Structure of Causal Spaces." *Proceedings of the Cambridge Philosophical Society*, 63,, pp. 481-501.
- Nicolis, G., and I. Prigogine. 1977. *Self-organization in nonequilibrium systems: from dissipative structures to order through fluctuations*. Wiley Toronto.
- Prigogine, I., and I. Stengers. 1984. *Order out of Chaos*. New York.
- Sjödin T. & Heylighen F. 1985: "Tachyons Imply the Existence of a Privileged Frame", *Lettere al Nuovo Cimento* 44, p. 617-623.
- Smolin L., Did the Universe evolve?, *Classical and Quantum Gravity*, 9(1) 1992, 173-191.
- Smolin, L. 1997. *The Life of the Cosmos*. Oxford University Press, USA.
- Stadler M. & Kruse P. 1990 *Theory of Gestalt and Self-organization*. In: Heylighen F., Rosseel E. & Demeyere F. (ed) *Self-Steering and Cognition in Complex Systems*. Gordon and Breach, New York, pp 142-169.
- Turchin, V. F. 1993. *The Cybernetic Ontology of Action*. *Kybernetes* 22, p. 10.
- von Foerster, H. 1960. On self-organizing systems and their environments. *Self-Organizing Systems: . M.C. Yovits and S. Cameron*. Pergamon Press, pp. 31–50.