

Andree C. Ehresmann, Jean-Paul Vanbremeersch

Knowledge in Memory Evolutive Systems

Abstract

The Memory Evolutive Systems (MES), introduced in a series of former papers, propose a mathematical model, based on Category theory, for open autonomous hierarchical systems such as biological, neural or social systems. Here the authors show that MES give a framework in which to study the formation, representation and interpretation of knowledge. In particular, they explain how the system can acquire pragmatic or conceptual knowledge through the coordinated action of a net of competitive internal patterns of agents, called coregulators (or CRs). This knowledge is not rigid, but is adapted to different situations subject to an appropriate choice of some parameters. Depending on its complexity level, a CR participates in an automatic or deliberate manner to the development of the collective distributed memory, relying on the partial information it collects at its own timescale on the system, its environment and the records of its past experiences. Higher CRs can classify records and develop more complex conceptual knowledge. Language (for men and social groups) allows a system to interpret its own knowledge and intentionally diffuse it.

1. Introduction

Let us give the following examples:

1. A spring keeps the «memory» of its shape; when it is pulled, it elongates but afterwards it comes back to its initial shape.
2. A thermostat measures the temperature and «knows» at which temperature I must stop the heating.
3. An expert system has knowledge with respect to a particular domain, and can answer questions on this domain.
4. A robot has sensorial organs to recognize some features of its environment, and several strategies (built-in or learnt) to react in an appropriate manner.

5. Bacteria have a metabolic activity, reproduce and are able to repair damaged DNA during the replication. These activities are autonomously controlled, using the genetic program.

6. A population of bacteria can adapt to changes in its environment (resistance to antibiotics), thanks to natural selection which favors bacteria with an appropriate mutation.

7. An animal with a nervous system gathers information on its environment by its sensorial organs, and on its internal state (hunger, pain, ...), and reacts by innate (holding reflex of a baby) or learnt behaviors.

8. Higher animals develop a semantics, generating a primary consciousness which may modulate their action depending on several parameters. And they exchange information through communication (alarm sounds, education of the youngsters, cultural differences among communities of chimpanzees (Whiten *et al.*, 1999)).

9. Human language leads to more efficient representations and communication; conceptual knowledge is developed and transmitted, generating culture.

a) *Several types of knowledge*

1. What is the common basis of these examples? Each of them exemplifies some kind of knowledge, in so far as this term is accepted with a large meaning covering the opposite pairs:

- cognition relying on natural processes of treatment of information (physical, biological, social, ...)/intentional knowledge emerging from it,
- innate (built in or inherited)/learnt,
- pragmatic (know-how, skills,...)/conceptual,
- automatic/deliberate,
- comprehensive/specialized,
- attributed to the agent by an external observer/explicit for the agent,
- distributed between several agents/controlled by a unique agent.

2. More explicitly: some knowledge is inherent to the system, either if it has been implemented in it by construction (examples 1, 2, 3, partially 4) or, for a living organism, if it is inherited or the consequence of natural selection (examples 5, 6). On the contrary, it can be acquired by learning, education or culture (examples 7, 8, 9).

In examples 4, 5, 7, the behavior depends on practical knowledge (built-in, innate or learnt strategy), to react in an appropriate manner to some external situations. It remains implicit for the agent, to which it is attributed only by an external observer. In the examples 8, 9 (and partially in examples 4 and 7) the agent controls intentionally part of its behavior, though some part remains hidden. We know how to pick up an object, but without knowing how our muscles do it.

In examples 8 and 9, there is also some conceptual knowledge, relying on a semantics; an intentional agent develops it combining simpler knowledge by several means, up to logical operations. For man, language helps the agent to interpret its own knowledge. It is communicated by imitation, education, directly or through some material support (book, file, CD, ...).

A particular agent has its own more or less specialized knowledge (examples 3, 4, 5, 7, 8), but comprehensive knowledge is generally distributed among the members of a group (examples 5, 8, 9), each one having only a partial access to it (no mathematician can grasp the whole of mathematics). Distributed knowledge can remain implicit for the group, such as the usual unconscious social comportment analyzed by Goffman (1973).

b) Formation and interpretation of knowledge

1. The knowledge of a (not necessarily living) system consists of internal representations in concrete relation with some features of the environment or some activities. They are innate, or constructed to memorize information received by the system under the form of an internal change (strengthening of a synapse, change of probability between various conducts, sign, record,...). The information will be retained only if it is stable enough (e.g., repeated several times) or significant: we continuously get new sensory information on the objects around us. but we discard most of them. Learning transforms a perceptual configuration more or less briefly activated by an external event into a stable internal representation; this «record» takes on its own identity by consolidation, and may be recalled later on by the same or by an approximate perceptual configuration.

2. Using approximate knowledge can help to react quickly to various situations, and later lead to a finer adaptation to them through the change of some parameters.

But this plasticity is counteracted by the risk of errors coming from insufficient information, an inadequate analysis of the situation (optical illusions), or modifications in the context. For example, in classical (Pavlov, 1927) or operant (Skinner 1938) conditioning, the conditions of the stimulus/response experiment can be changed, confusing the animal. For conceptual knowledge, errors can

also result from a wrong interpretation of the concepts used; most errors done by students in Mathematics are of this type: they interpret a mathematical concept (say, a derivative) from a curtailed representation such as a learnt formula which has not been integrated as an object on which other processes can operate.

5. Memorized data will be called «knowledge» only if they are interpreted as such either by the «knowing» agent or by an external observer able to attribute this knowledge to the agent. Thus knowledge is a ternary relation: knowledge of

something attributed to some agent by an interpreter (possibly the knowing agent itself). Here agent and interpreter can be living organisms, groups or machines.

The attribution of knowledge to the agent by the (internal or external) interpreter is based on:

- direct observation of the behavior,
- (partial) reading of the memory of the agent,
- material traces produced by the agent (books, files, CD, ...),
- inquiry to the agent (second degree: the agent must already interpret his own knowledge).

In each case the attribution can be false since the interpreter has only an external and partial view of the agent's memory (even if it is a conscious agent itself). For instance, the principle of charity (Quine 1960, Dennett 1990) attributes an action to a rational compoment; but the agent can act for other reasons unknown to the interpreter, or voluntarily induce him into error. Two different interpreters (e.g., one being the agent itself) can attribute different knowledge to the same agent: a teacher can judge that a pupil does not know the lesson which the pupil thought he knew.

Moreover the interpretation always depends on a specific context, because knowledge evolves. For instance, ancient Greeks «knew» that the earth is flat, though it is not. Even knowledge relying on a consensual authority and obtained by accurate methods, say the scientific method, can be falsified later on; think of the changes of paradigms in science (Kuhn, 1972), or the different perception of past history depending on the context.

c) Memory Evolutive Systems

1. For a system to have knowledge, it must at least have some of the following capacities:
 - to gather information via modules linked to the exterior (receptors), to respond by adapted strategies (effectors), and to evaluate the result of these strategies, at least locally (locate fractures in some modules),
 - to record information which are stable enough or repeat themselves, as well as its strategies and their result, so that it develops a memory formed by their internal representations and can modify it to account for its successive experiences,
 - to recall records in the memory, connect them and operate on them, e.g. to learn new strategies formed by combining already known strategies,
 - to interpret its knowledge, or to be observed by an external interpreter,
 - possibly to classify the known objects according to a semantics; this allows for conceptual knowledge and its development by logical processes

(disjunction, conjunction, negation), by induction (generalization), deduction (formation of chains of links, as in mathematical proofs), or abduction (find the causes of a situation),

- to communicate with other systems (e.g. members of a same group) to develop common knowledge, up to culture.

2. A model of such systems will be given by the *Memory Evolutive Systems* (MES) which we have developed since 15 years; it is a mathematical model, based on the theory of categories introduced by Eilenberg and Mac Lane in 1945 (cf. Mac Lane 1970 for the main definitions). In Section 2, we describe this model in a concrete setting, and refer to our preceding papers (a list is given in our Internet site, Ehresmann & Vanbremeersch 1999) for a more theoretical approach, which is only alluded to here (in paragraphs beginning by #). Section 3 analyses the processes of acquisition, consolidation and cohesion of knowledge, distinguishing the different types of knowledge. In Section 4, we show how more complex systems may develop a conceptual knowledge and interpret it, leading to its diffusion on a larger scale.

2. Organization and functioning of a system with knowledge

Here we describe the organization and functioning of a system in which knowledge plays a role. The system can be a living or artificial organism, a neural system, an animal group (e.g. a hive), a human society, ... It will be modeled by a Memory Evolutive System (MES).

a) *Description of the system*

1. The system has a given *timescale*, from its «birth» to its «death».

Its state at a time t of its life is determined by its present organization consisting in:

- Its components at this time, among which we distinguish: (i) its *agents* which are its individual constituents (neurons, bees, members) and also more or less complex associations of these constituents participating in common activities (visual areas in a neural system, the class of workers in a hive, departments of an enterprise); (ii) objects constituting its memory, also called *records*, they are internal representations of the knowledge of the system, which encompasses known features of the environment, strategies to deal with at any level (from metabolic regulations up to complex skills), possible conceptual knowledge and its material support; (iii) more or less temporary objects, acting as information, which are the internal traces of the signals received by the system from the environment at t .

- Relationships (called *links*) between these components which allow the implicit or explicit transfers of data, energy, constraints between agents, the reception of information, its possible recognition by recall from the memory or its later storage, and commands of adapted strategies to effectors. Each link operates with a specific delay of propagation.
- Successive links can be composed, and the composites of two paths of links are identified if they are functionally equivalent.

The agents, their knowledge, the information they receive and their interactions change in time. The change is measured by the *transition* from the state at t to the state at a later time t' , which indicates which of the components and links existing at t are still there and what they have become at t' , as we could recognize a particular member of a group on two successive photographs of the group. Thus a component of the system, say an agent N , is not represented by a unique invariant object (as in usual models), but by the sequence (N_i) of its successive states all along its life. And the same for the links between components.

We have so modeled the system by an *Evolutive System* (Ehresmann & Vanbreemers 1987), which is defined by:

- a (finite or infinite) part of the real numbers \mathbb{R} representing its *timescale*,
- for each time t , a category¹ K_t , the *state-category* at t ,
- for each time $t' > t$, a functor² $k_{\{t,t'\}}$ (the *transition* from t to t') from a subcategory of K_t to $K_{t'}$, these transitions being transitive, i.e.,
if $k_{(t,t')}(N) = N_t$ is defined and $t' < t''$, then $k_{(t',t'')}(N_{t'})$ is defined iff $k_{(t,t'')}(N_t)$ is defined, and then both are equal.

2. All the components are not of the same complexity level. An agent which is an association of other agents (a department of an enterprise) is more «complex» than these agents (its employees), and it can itself be one of the constituents of a more complex agent (division regrouping several departments). In the memory, a complex skill is formed by coordinating together more elementary skills. Even a cell has a whole hierarchy of lower level components from its atoms up (Chandler 1997).

¹ A(n oriented multi-)graph consists of a set of vertices N and a set of arrows between them. A *category* is a graph on which there is given a composition law associating to each pair of successive arrows *if* $N \rightarrow N'$, $g: N' \rightarrow N''$) an arrow $fg: N \rightarrow N''$; this law is associative (a path has a unique composite whatever its 2-2 decomposition), and each vertex N of the graph has an «identity» arrow $id_N: N \rightarrow N$ whose composite with any arrow h beginning or ending at N is equal to h . The vertices of the graph are called the *objects* of the category, the arrows its morphisms, or more concretely here, its *links*.

² A *functor* from a category to another is a mapping respecting their graph structures and their composition laws.

Thus the system has a hierarchical structure: a component N of a given level is obtained by binding together a pattern formed by components of the next lower level with some distinguished links between them; such a decomposition pattern represents a (not necessarily unique) internal organization of N .

The links between two components N and N' can be «simple» in the sense that they are obtained by binding a compatible family of links (or «cluster») between the lower level components of N and N' . But there are also «complex» links which emerge at higher levels. In Section 3, we'll come back on the genesis of these links and their role in the interconnection of all knowledge.

The Evolutive System is *hierarchical*, in the sense (Ehresmann & Vanbremeersch 1987) that its components are divided into several complexity levels, with a component of level $n+1$ being the *colimit of a pattern?* (or inductive limit of a diagram in the sense of Kan 1958) of linked components of level n .

And it has a hierarchical sub-system, forming its *memory*.

3. The changes are essentially attributable to what Thom (1988) calls the archetypal operations: „birth, death, scission, collusion”. Some agents will disappear while others arrive. Some signals will be memorized and become knowledge, while others will be discarded. New complex components are formed: agents can assemble to form a new group having some competence, more complex skills are learnt,... Conversely, a sub-group can dissolve.

The transitions between the state-categories are constructed by the process *complexification with respect to a strategy* (Ehresmann & Vanbremeersch 1987). A *strategy* on Kt consists of a set of external elements A to add, a set of patterns P to bind together, a set of more or less complex components B to suppress. The complexification is a new category which is explicitly constructed (as a special case of the construction of the prototype of a sketch given by A. & C. Ehresmann in 1972):

- Its objects are: those of Kt except the B 's, the added elements A , and, for each pattern P to bind, a new higher level object CP which becomes its colimit.
- There are two kinds of links from CP to a CP' . The (RP') -*simple links* bind clusters from P to P' . A *cluster* is a maximal family of links from the objects P . of P to those of P' satisfying the condition: there is at least one link from each P . to some object of P' , and if there are several they are correlated by a zig-zag of links in P' .

A *pattern* (or diagram) P in a category K consists of a family (P_i) of objects of K and some links x between them, called its distinguished links. A *collective link* from P to an object N of K is a family of links $f: P_i \rightarrow N$ well correlated by the distinguished links of P , i.e. such that $x_i f_i = f_j$ if for each $x: P_i \rightarrow P_j$. A *colimit* of P is an object C of K such that the links from C to any object N of K are in a 1-1 correspondence with the collective links from P to N .

- The *complex links* from P to P' are obtained by composing a sequence of simple links binding non-adjacent clusters.

b) Local regulations

1. The preceding description is purely formal and «external»: it could be given only by an observer with a complete view of the system, and that cannot exist for complex enough systems, specially autonomous systems (Matsuno 1989). In particular knowledge is distributed among the agents; it is constructed through their combined action and later on takes its own identity. For instance, the bees of a hive «know» how to construct the hive, but each one participates in a very fragmentary way to this process, thanks to some instinctive strategies; the «construction of the hive», which is a consequence of the temporal combination of all these strategies, figures among the knowledge of the system, but it remains hidden to the bees and can be attributed to their society only by an external observer.

2. The system is autonomous in the sense that it is internally controlled by the agents. While some agents cooperate, there are also competitions or even conflicts between them. We call a *CoRegulator* (CR) a subsystem formed by a small pattern of agents of the same complexity level (possibly forming a higher level agent), acting together at a specific discrete timescale (e.g., an operon in the genome of bacteria). The global dynamics is modulated by the competition between a net of CRs which operate in parallel, but with different rhythms; e.g., in an industry, a workshop has a daily cycle, while design departments can plan over several years.

With respect to knowledge, each CR has a partial and differential access to the collective memory which it contributes to develop by operating a stepwise trial-and-error learning process at its own timescale.

The system is thus modeled by a *Memory Evolutive System* (MES): it is a hierarchical evolutive system, with a hierarchical sub-system called the *memory*, and a net of evolutive sub-systems with discrete timescales, its *CRs* (Ehresmann & Vanbremeersch, 1991).

3. One step of a particular CR extends between two successive dates of its timescale; it is divided in several more or less overlapping phases (forming an epistemo-praxeological loop in the sense of Vallee 1995), which we illustrate by a meeting of the editorial board of a Journal:

- In the first phase (or *actual present*), formation of the actual *landscape* of the CR, which is a (more or less distorted) internal representation of the system for the CR; it filters the partial information received by the agents during their actual present, and plays the part of a working memory during

the step. For instance, the editors will register the various papers submitted since their last meeting, the referees' reports on papers formerly received, the letters sent by readers, and control if the decisions taken at their preceding meeting have been correctly carried out.

- Selection on the landscape of a strategy to react in an adapted way to the context, to memorize the stable enough information, possibly combine them into new knowledge and/or transmit them; the choice is supported by the recall from the memory of former similar situations. The board will select the papers to be published in the next issue of the Journal, the referees for newly received papers, the intended schedules; for this, they will recall preceding reactions of readers, which referees have previously done a good job, and former delays of the printer. The strategy can be really «chosen» by the CR (intentional action), or imposed on it by other CRs (e.g., the direction of the Journal), by external constraints (excessive cost for too long papers), or represent a known «automatic» answer to the given situation (always the same referee for some kind of papers).
- Commands to effectors to realize the strategy: the papers for the next issue are sent to the printer, the new ones to the chosen referees. The landscape gives only a partial and a more or less flawed view of the system and the various CRs may conflict, so the objectives of the strategy are not necessarily fulfilled. The step can even be interrupted by a *fracture* if no strategy can be found (the editors cannot agree on which paper to publish), or if the selected strategy cannot be effected (the printer refuses to continue printing the Journal). Such a fracture may reveal a lack of internal coherence in the knowledge of the system, or a wrong correspondence with «Reality» (whence the problem of Truth...).
- If the step processes without a fracture, at the next step, the result is evaluated by comparing the anticipated landscape with the newly obtained landscape; and the strategy is memorized with its result. At their next meeting, the editors will verify if the issue is well printed and the reports of referees received. They will note if the printing delays have been respected and no complaint has arrived.

The *actual landscape* L at t is the category whose objects are the perspectives for the agents of the CR of the components B of the system of a near complexity level; a *perspective* of B is a cluster of links from B to the pattern formed by the CR during its actual present. There is a *distortion* functor from L to the system. The anticipated landscape for the end of the step should be the complexification of L with respect to the selected strategy, and it is compared to the actual landscape effectively obtained at the next step by a *comparison* functor.

c) *Global dynamics*

1. The strategies selected by the various CRs at a given time are not realized on their landscapes but relayed to the system where they are not always compatible. Indeed, the CRs share common resources, have differing perspectives on the distributed knowledge and there are direct and indirect interactions between them, so that conflicts may occur between their strategies. The editorial policy may clash with the economic constraints of the publishers of the Journal.

An equilibration process will arise between the strategies, called the *interplay among the strategies of the various CRs*. It is not a centrally controlled process, but a dynamical modulation between the different relayed strategies. It depends on the respective «weights» of the strategies and of the CRs (the editors can argue with the publishers). A main role is played by the structural temporal constraints of the CR: a paper or report not received at the date of the meeting cannot be examined, the printing of the issue can be delayed if the printer has too many other works to do.

If the constraints of a CR cannot be satisfied, a fracture occurs in its landscape and, if it is not quickly repaired, there is a dyschrony: the regular publication schedule cannot be resumed before several issues. However fractures can have a creative role, by imposing a complete overview of the situation. If the printer cannot respect the delays, a new printer can be chosen, perhaps making also a better job.

The *structural temporal constraints* of a CR at t connect its *period* (mean length of a step) $d(t)$ to the mean propagation delays $u(t)$ of the information it receives and to the mean stability spans (cf. section 3, b) $v(t)$ of the components intervening in its landscape and the strategy: For almost all t (i.e. except on a set of measure 0), we must have:

$$u(t) \ll d(t) \ll v(t).$$

2. Thus a *dialectics* via functional loops is generated between CRs which are *heterogeneous* with respect to their complexity level and/or their period. A series of fast changes by lower CRs is only perceived as a whole and with a delay by a higher CR, and may cause a fracture in it; to repair the fracture this CR may impose new strategies on the lower CRs, and the process goes on. If the editors progressively modify the contents of the Journal, the publishers will perceive the change only after a delay, but then it can displease them, and they may react by dismissing some editors.

This dialectics shapes the evolution of the system and differentiates it from «simple» physical systems (Rosen 1986). It leads to the formation of more and

more complex strategies and to the development of a coherent corpus of knowledge. For instance, it explains how culture can be transmitted, or temporally refrained, and how it affects the comportment of the individuals who receive it.

Let us give a personal example of development of a mathematical theory illustrating this dialectics. In the late sixties, a small group of young research students working with Charles and Andree Ehresmann has developed the theory of sketches; they had frequent mutual exchanges, so that they had adopted particular concepts and even notations. But they had almost no contacts with the main stream of categoricians, in particular in the States, and their work, not well published, remained unknown. At the first conference in 1970 where their results were exposed, the «establishment» could not understand them, because they were far from the current problems (topos and triple theory) and the notations were unusual. This cold reception caused a fracture to the young students. But it had also a beneficial effect: contacts were established, specially thanks to the organization of „Jounees Theorie et Application des Categories” in Paris and Amiens, and of international conferences in Amiens in 1973, 1975 and 1980. These meetings allowed to harmonize the notations and better explain the motivations. The consequence has been a diffusion of the theory of sketches, which has been-widely adopted in the eighties, with important applications in Computer Science (Barr & Wells 1984, Gray 1989, Walters 1991).

3. Development and plasticity of knowledge

The knowledge of the system is represented by the content of its memory. It is based on a kernel of innate knowledge which is later developed by learning. Its role is essential to recognize objects or configurations already met and to respond in a more adapted way.

a) Acquisition of knowledge

One of the objectives of the system will be the formation of records and links to memorize new situations and strategies. It will be achieved through the strategies of the CRs and the interplay among them.

1. A new configuration C met in the environment (say, an unknown object) or produced by the system (e.g., commands of a new strategy) is internally represented by the temporal coordination of a pattern P of components; its objects can be receptor agents, or components linked to them such as records in the memory recognizing parts of C. The configuration will be memorized by the formation of a new object of the memory, called the *record* of C, which integrates

the pattern in a higher level unit by strengthening its links. For this, each CR will memorize the attributes of P which it may distinguish, and the interplay among their strategies conjugates their actions to form the record M . A later occurrence of C will reactivate the strengthened pattern P (which we call a *decomposition* of M), thus leading to the recall of M and the recognition of C .

For instance, let us analyze how an orchestra will learn a new partition C . Each player acts as a particular CR. A pianist E sorts out the part for piano, which he translates in his actual landscape into a sequence of notes and tunes to be integrated as a unit in his own memory, called the E -record of C . Independently the other players learn their part. During the repetitions of the orchestra, the interplay among the (strategies of) the players will synchronize and harmonize their parts, and integrate them into an object M of the collective memory of the orchestra, which we call the *record* of C ; the various E -records of C become a reflection of this global record. If the orchestra takes back the partition later on, each player will easily recall his part by reactivating his E -record, and they'll have no problem synchronizing their parts to reactivate the global record.

A new configuration C is internally represented by the formation of a pattern P of linked components synchronously activated. A particular CR, say E , will perceive in its actual landscape a pattern of perspectives coming from a sub-pattern (possibly void) P_E of P . An objective of the strategy of E will be to memorize this pattern p_E . This will be reflected to the system into the command to bind together the pattern P_E (image of p_E by the distortion functor). Simultaneously, other CRs form their own record of C . The interplay among the strategies will integrate the various commands, so that the complexification process with respect to the global strategy thus obtained will add a colimit M of P , called the *record* M of C ; its perspective in the landscape of E becomes a colimit of p_E , called the *E -record* of C . A later presentation of C reactivates P , hence also its different E -records which, by the interplay among strategies, are binded together, thus the recall of the record M .

2. The formation of the record of C corresponds to its „assimilation” in the sense of Piaget (1940). Afterwards it will be consolidated and adapted to gradual enough temporal modifications of the system and of the environment (Piaget's „accommodation”). In this way, the record takes its own identity as a component of the system, with its successive states becoming more and more independent from the particular decomposition used in its formation.

For instance, the replacement of one or two players will not prevent the orchestra to play C , with minor variations of its record; a law can be revised to keep track of progressive changes in a society; a scientific theory can be adapted to new facts, before a change of paradigm (Kuhn 1972). The rate of change is measured by the stability span of the record; during stability periods, the span is long, while it is shorter during periods of development or of decline.

If M has been constructed at t as the colimit of a pattern P of lower level components, the evolutions of M and of P may remain correlated during a certain period, but diverge at a time $t' > t$, so that the state of M at t' is no more the colimit of the new state of the pattern P (e.g., if objects of P have been suppressed or replaced). The *stability span* of M at t is defined as the larger real dt such that there exists a pattern Q of lower level components whose state, for each time s between t and $t+dt$, has for colimit the state of M at s .

3. The development of the memory does not only consist in the formation of records, but also in the formation and strengthening of links between them, thus increasing the cohesion of knowledge. Let us mention some of these links.

When a configuration C is memorized, the links between the objects P. of the internal pattern P it activates are strengthened; and new links are formed from each P to the record M of C (heredity links). Later M may participate to the formation of the record of a more complex configuration having C as a constituent, and thus become linked to this new record.

A link can also be formed from the record M to (the record of) a strategy in response to C; e.g., classical or operant conditioning (Pavlov 1927, Skinner 1938) creates such links between a stimulus C (the sound of a bell) and a conditioned response (the dog salivates). A later presentation of C recalls M, and, if the link is strong enough, it will-lead to the automatic recall of the strategy. This is done, as above, via the different CRs and the interplay among their strategies, with a risk of fracture if some structural temporal constraints cannot be respected.

For instance, the view of a prey has no effect for a satiated animal; but if he is hungry, he'll try to catch it. The catching strategy requires a coordination between the visual CRs which determine the location and size of the prey, and the motor CRs which control the motion of the predator. If the prey runs too fast, or in an erratic manner, the visual information on its location arrive too late to the motor CRs to adjust the movement, and the prey flees.

The memory is an evolutive sub-system of the MES; its transitions correspond to complexification processes with respect to strategies whose objectives are the formation of new records. The construction of a complexification explicitly determines which links are formed between the added records (cf. Section 2, a).

b) Plasticity of knowledge

1. We have seen how a record M is consolidated to adapt to small temporal variations of the context. But there is a more comprehensive kind of consolidation, by extension of its domain of application.

In particular, if a strategy has been successful in some situation and if a near enough situation arises, the different CRs will try to use the same strategy. If it succeeds, its domain is extended; if it fails and causes a fracture to some CRs, the fracture will be repaired by modifying the corresponding sub-processes to adapt it to the new context. For instance, a dog learns to bring back a particular ball sent by his master in the garden; later on, he will use the same strategy to bring back another object sent by someone else in an other place.

A record adapted to several contexts will be called *multifold*, it will be recalled by the activation of a specific decomposition P in each context; such a decomposition can be thought as the fixation of the values of some parameters. As we will see, the fact that records can be(come) multifold explains the development of complex relationships between them, leading to an interaction of all knowledge.

In a MES, a component M which has been formed as the colimit of a pattern P may also be, or later become, the colimit of patterns Q non equivalent to P (Multiplicity Principle, Ehresmann & Vanbremeersch 1996); we then speak of a multifold component, and the passage from P to Q is called a *complex switch*.

2. There are two kinds of links which are formed between records: the simple links and the complex links. For instance, if M is a text written in French and M' its English translation, we have a simple link «translation» from M to M' associating to each sentence in M its translation in M' . But the link between a sentence in M and its translation in M' will generally not be the simple translation of each word, because of the differing structures of the two languages (whence the difficulty of an automatic translation!).

More generally let M and M' be two records. If P is a decomposition of M and P' a decomposition of M' , a (P,P') -simple *link* from M to M' binds together a cluster of links between the objects of P and P' .

By composing a chain a simple links, we get a link, but it is not necessarily simple. Indeed, if M' is multifold, it can be recalled in another context through a decomposition Q' non-equivalent to P' (corresponding to a different choice of parameters). Then if we compose a (P,P') -simple link from M to M' with a (Q',P'') -simple link from M' to a record M'' , the ensuing link from M to M'' may not be (P,P'') -simple; we then say that it is a *complex link*.

A mathematical example illustrates the difference between simple and complex links. A topological space is the geometric realization of several simplicial complexes. If P and P' are two simplicial complexes associated to the topological spaces T and T' , a (P,P') -simple link from T to T' is reduced to a simplicial map from P to P' ; but a complex link from T to T' is any continuous function.

The development of the memory comes from a sequence of complexifications. Each one introduces simple links and complex links (cf. Section 2, a).

c) Cohesion and complexification of knowledge

The existence of complex links is at the root of the development of more and more complex and intricate knowledge.

1. Complex links establish comprehensive relationships between records. Indeed, a complex link from M to M'' relates not only these two records, but also the intermediate multifold records which occur in its formation through switches between two of their decompositions (change of parameters). Thus the cohesion it creates between M and M'' reflects more than a «local» cohesion between lower level decompositions of M and M'' ; it reflects something of the overall structure of the lower level memory (containing the decompositions), emerging at the level of the link.

For instance, chains of inferences using metaphors (Paton 1997) are powerful to reveal new overall outlooks because they correspond to the formation of complex links; indeed, a metaphor can be interpreted as a switch between two recompositions of a same record: the genome with its chemical structure, or looked at as a text.

If a record has a decomposition with some complex links, it inherits not only local properties from the objects of the decomposition, but also new comprehensive properties which emerge through these complex links and rely on implicit assumptions. This cohesion of knowledge increases the risk of ambiguities in communication between people or systems having different kinds of knowledge, for they may not share the same implicit knowledge. An «expert» may have problems to disentangle all the data necessary to construct an efficient expert system. Robots behave well only in a very simple environment, where all the conditions can be controlled.

If a complex link gg' is the composite of a (P, P') -simple link g from M to M' with a (Q', P'') -simple link g' from M' to M'' , its properties are deduced from the «local» properties of the two clusters, say of level n , that g and g' bind, but also from the fact that P' et Q' have the same colimit M' . This last condition implicates the global structure of level n (before the complexification), since the «universal» property of the colimit M' means that the two patterns impose the same constraints to any object.

2. The memory is hierarchical; when records of a given level have been consolidated and connected by simple and complex links, they can be assembled to form more complex records (or hyperstructures, Baas 1992), by iteration of the preceding formation process. A higher level record A in the memory necessitates several stages to be formed (to compare with the „stades” of Piaget 1940); the number of these stages characterizes the *complexity order* of A .

For instance, in a 2 stages process, first the strategies of the CRs and the interplay among them single out and coordinate patterns RZ of existing records with (simple or complex) links between them, to form new records of a higher level. After their consolidation, the second stage similarly sorts out a pattern R with these new records, and memorizes it by the formation of the record A; we say that $(R, (RZ))$ is a *ramification* of A of length 2. In the following consolidation process, A may acquire other ramifications.

Later on, A can be recalled through the unfolding of anyone of its ramifications. This consists in the synchronous activation of the objects of a decomposition of A, through different CRs (in their actual landscapes); and, at a following step, the same process, applied to each of these objects, activates one of its decompositions, through lower level CRs, and so on if there are more stages. The choice of the ramification is done step by step, from top to bottom: first choice of a decomposition of A, then, for each of its objects, choice of one of its decompositions, ... Unfolding of a ramification can be compared to the gradual filling of the different slots of a frame in the sense of Minsky (1986). Which ramification is finally unfolded will depend on the context, and it is selected at each step through the interplay among the strategies of the CRs, taking into account their structural temporal constraints. If these cannot be respected, the action will fail.

For instance, before he may walk, a child learns to coordinate more or less innate strategies to stand up and to move forwards a leg when he is held; the corresponding strategies are memorized in the motor areas of his brain. And these strategies are themselves coordinated into patterns allowing to make a step without falling. And so on up to the formation of a complete strategy for walking, usable in the most varied situations (Josephson, 1998). But the coordination must always respect temporal constraints; if the child tries to walk too quickly, he falls.

A is constructed as an *iterated colimit*, i.e., A is the colimit of a pattern

R of objects R_i in the memory such that each R. is itself the colimit of a pattern

R_i , such that each object R_{ik} of R_i is the colimit of a pattern R_{ik} , and so on.

We then say that A is the iterated colimit of the *ramification* $(R, (R_i), (R_{ik}) \dots)$.

Iterated colimits are formed through successive complexifications of the memory. It follows from the Multiplicity Principle that A may also be (come) the iterated colimit of other ramifications.

The *complexity order* of A is the smallest length of a ramification of A down to the lowest level; it is less than, or equal to, the level of A. In particular, if A is the colimit of a pattern of objects of order n , we have proved that A can also be of order n if all the distinguished links of the pattern are simple, but its order is $n+1$ if at least one of these links is n -complex (Ehresmann & Vanbremeersch 1996).

d) Mental objects and higher order cognitive processes

1. In a neural system, the record of a simple stimulus is reduced to an individual, specialized neuron; for instance, in the visual area, there exist „simple cells” representing a segment of a given direction, and „complex cells” representing a particular angle (Hubel & Wiesel 1962). But more complex stimuli, except for some exceptions (e.g., a neuron activated by a hand holding a banana for a monkey, Gazzaniga 1985), do not have their own „grand-mother neuron”.

The development of neural imaging shows that complex perceptual stimuli, or motor programs, are represented by the short-lived synchronized firing of a specific assembly of neurons. And learning would consist in the formation of such synchronous assemblies, through the strengthening of synapses between their neurons, following the rule already proposed by Hebb (1949): a synapse between two neurons is strengthened if the two neurons fire at the same time, and depressed if one fires while the other does not.

In the MES modeling the neural system, the record corresponding to such a synchronous assembly of neurons is called a *category-neuron* (or briefly *cat-neuron*); it operates as a unique «higher order neuron» integrating the synchronous assembly. More complex records, also called *cat-neurons*, are constructed in several stages. *Cat-neurons* of order 2 correspond to a super-assembly (or «assembly of assemblies») of neurons, which cannot be reduced to a (even large) synchronous assembly of simple neurons; higher order *cat-neurons* correspond to synchronous super-super-assemblies, and so on. They represent higher order mental objects and cognitive processes.

The above description of the links between records determines what are the possible interactions between synchronous (super-)assemblies of neurons, thus solving a problem raised by neuroscientists (von der Malsburg & Bienenstock 1986), which cannot be approached by classical models: they are the simple and complex links between the corresponding *cat-neurons*. And it becomes possible to «compute» with *cat-neurons*, i.e., with (super-)assemblies of neurons, as if they were simple neurons, thus developing a real „algebra of mental objects” following the proposition of Changeux 1983).

Let us remark that this model is very different from neo-connectionist models of neural systems which give only a description at the sub-symbolic level, and for a limited period, without taking into account the interactions between the different levels. In particular, these models can only describe the formation of simple *cat-neurons* (represented by attractors of the dynamics), but not their complex links, so that they cannot describe higher order *cat-neurons* modeling complex mental objects.

In the MES modeling a neural system, the state-categories are obtained by successive complexifications of the *category of neurons* defined as follows: the objects are the neurons, the links between them are polysynaptic pathways; two polysynaptic paths from N to N' are identified if they have the same strength, i.e., if the probability that they propagate the activity of N (determined by its instantaneous frequency of spikes) to N' is the same, as well as the delay of propagation.

2. The representation of higher order processes by cat-neurons leads to a new approach of the philosophical problem of the identity between mental states and physical states of the brain. Indeed, a physical state, as it is seen through brain imagery, corresponds only to the activation of a synchronous assembly of neurons, modeled by a «simple» cat-neuron. But a mental state is represented by a higher order cat-neuron whose activation requires a several steps unfolding through the various intermediate levels of a ramification, down to the level of physical states; and at each step, it can proceed along one or another non-equivalent decomposition of multifold records, with possibly a switch between them, the origin of which might be random (neural «noise»), quantic (as proposed by Eccles, 1986), or controlled. Though such a process represents a well described «physical event», we cannot identify it with a «physical state»: mental states emerge in a dynamic way (through the gradual unfolding of a ramification) from physical states but are not identical to. them. This could qualify as an *emergentist monism* in the sense of Bunge (1979).

4. Classification and interpretation of knowledge

The knowledge of a system is distributed among its agents though it remains mostly implicit for them, and they have only the capacity to use part of it in appropriate situations. However there may exist higher CRs which interpret and classify the automatic knowledge of lower level CRs, and develop a semantics allowing for a more flexible and deliberate application of knowledge. Here we just delineate the main ideas, referring for more details to the article on consciousness in our Internet site (1999).

a) Semantic memory

The consolidation of a record consists in its adaptation to several circumstances by a change of parameters (unfolding of different decompositions and ramifications). These parameters can come from the recognition of constancy through variable circumstances, so that a same response can be given to a whole

class of similar situations; a frog will jump after a fly or after any flying object of the approximate size. This relies on a classification of the records contained in the memory into what we call «concepts».

1. A lower CR, say E, operates only an automatic classification of the records according to the attributes it distinguishes: two records are «acted» as equivalent if their traces in its landscape activate the same pattern of agents. For instance, birds recognize squares, whatever their size, from triangles; or, at a more biological level, it is the same pattern of neurons of a color-CR which is activated by all blue objects.

But this classification remains implicit at the level of E, and it is interpreted as such only by a higher level CR, with a longer period, able to overview what is common to the different records in a same class. A pigeon does not «know» that he classifies geometric forms; this «knowledge» is attributed to him by the observer. This higher CR memorizes each invariance class under the form of an object, called an *E-concept*, whose *instances* are the different records of the class. For example, the instances of the color-concept «blue» are all the blue objects.

The CR-concepts, with respect to the various CRs, form a *semantic memory*, which is extended by the formation, firstly, of concepts simultaneously classifying several features (as a blue triangle), and then of more abstract *concepts* obtained by assembling such «concrete» concepts. A concept can be thought of as an abstract prototype for a class of objects with a „family resemblance” (in the sense of Wittgenstein 1953); it does not necessitate the existence of a language.

The automatic classification operated by a CR, say E, is modeled using the notion of „same shape”, in the sense of Borsuk (1975) „shape theory” (cf. Cordier & Porter 1989). An E-concept will be defined as the (projective) limit of the pattern of agents of E activated by all the instances of a given invariance class. More general concepts are obtained as iterated limits of patterns of E-concepts (for different E) linked by complex links; they are formed through successive complexifications.

2. The development of semantics increases the plasticity, but also the fine-tuning, of knowledge. Indeed, if a record M has a decomposition P, a new decomposition can be obtained by exchanging an object P of P by another instance of the concept of P.; the dog who has learnt to fetch a particular ball will later bring back any ball. Thus the number of ramifications of a record increases, making it useful in more contexts. Conversely, there is a fine-tuning of the classification: if the exchange of P. by another instance of the same concept does not function (causing a fracture), the classification will have to be refined; a child begins to assimilate all moving vehicles, but later learns to distinguish a car from a train. Thus extensive consolidation of knowledge, specially of

pragmatic knowledge, and development of a fine-tune semantics are interactive with mutual benefit, thanks to the recognition of distinctions as well as of similarities.

Moreover, the semantics leads to the formation of conceptual knowledge, based on records which are themselves concepts. In particular, the choice of a strategy by a higher CR will be done under the form of a concept, instead of a specific strategy of its invariance class. It adds a new degree of freedom in the interplay among the strategies of the CRs, since it will be possible to select among the strategies of the invariance class chosen by a CR the best adapted one, taking into account the strategies relayed by the other CRs.

b) *Intentional and conscious CRs*

1. A system can have, or develop, higher level agents with particular capacities for handling knowledge in a more deliberate manner. In MES, it is modeled by the existence and/or formation of higher CRs, which act as associative CRs controlling lower CRs.

A CR will be called *D-intentional*, or intentional in the sense of Dennett (1990), if it acts «as if» it was able to optimize its choice of strategy in its landscape. This is possible if some of its agents are evaluators which classify the records of strategies available in the landscape, depending on their result for the CR. However this classification can be «automatic», as in the case of reflexes. If it is deliberate, and the CR has access to the semantics, we speak of an *intentional* CR. It may be difficult to recognize if the behavior of an animal is intentional or only «as if».

2. An intentional CR is said to be *conscious* if it is able to internalize the semantics and the notion of time and eventually to interpret itself its knowledge. We characterize a conscious CR by the following capacities:

- Extension of its actual landscape by retrospection to lower levels of the near past, by an increase of awareness, in particular after a fracture which has revealed some errors in its handling of a situation, or lack of coherence in its knowledge;
- Development of an abduction process on this extended landscape to search for the possible causes of the fracture, taking into account preceding experiences;
- Complex programming of a selection of strategies covering several steps ahead, thanks to the formation of virtual landscapes in which sequences of strategies (selected under the form of concepts) can be «tried» without material cost for the system; these strategies can be recalled from the memory, or obtained by consolidation of already known strategies, or constructed by

assemblage of known strategies whose consequences are anticipated

(deduction). For instance, a chimpanzee solves the problem to seize a banana out of reach by putting a box on top of another and climbing on it.

3. There are several degrees of consciousness (cf. Edelman 1989, Damasio 1999), from primary consciousness up to self-consciousness. In a neural system, they rely on the existence of functional loops between various areas of the cortex, which form what Edelman calls a „loop of consciousness”.

Higher conscious CRs can interpret their knowledge and make a deliberate choice of long term strategies; however they cannot control all the consequences of these, and must rely on lower CRs to realize them; I decide to walk, but I don't know how to directly control the motion of my legs, whence the risk of falling if I try to think how it is directed.

Language, associating a word to a concept, permits the formation of more and more abstract conceptual knowledge, controlled by still higher conscious CRs which are able not only to interpret their knowledge but also to «know that they know». These CRs can communicate their knowledge by teaching, and by the production of material records. Whence the development and diffusion of culture.

5. Conclusion

The MES give a mathematical model, based on category theory, for complex autonomous systems. Their framework seems well adapted to study the problems related to the acquisition, representation, cohesion and interpretation of knowledge.

- Innate and acquired knowledge of the system is recorded in a collective distributed memory, where representations of features of the environment, past experiences and adapted behaviors are stored in a flexible way; their records are regularly up-dated to take into account new data, and their domain of application is gradually extended and refined.
- The dynamics of the system, leading to the processes of acquisition and consolidation of knowledge, is directed by the cooperation and/or competition between groups of agents acting as coregulators (the CRs); each CR has only access to part of the knowledge accumulated in the memory; and, depending on its complexity level, it more or less participates to the development of this memory through a «local» trial-and-error learning process;
- Knowledge remains pragmatic and mostly implicit for lower level agents; but higher level CRs in sufficiently complex systems (higher animals, social groups) are able to classify known objects into invariance classes, thus developing a semantics which allows for a deliberate and more precise tuning and handling of both pragmatic and conceptual knowledge.

- Agents endowed with «consciousness» can perform deduction and abduction processes, allowing for a better long term regulation of their comporment. If they possess language, they can self-interpret their knowledge and extensively communicate it, thus leading to the development of culture. *Acknowledgments:* We are grateful to Profs. Jerry Chandler and Brian Josephson for stimulating exchanges on related problems.

References

- Baas, N. A. (1992), Hyperstructures - a framework for emergence, hierarchies and complexity, in *Proceedings du Congres sur l'emergence dans les modeles de la cognition*, Telecom. Paris, 67-93.
- Barr, M. & Wells, C. (1984), *Toposes, triples and theories*, New York, Springer. Borsuk, K. (1975), *Theory of shape*, Warsów, Monografie Mat. 59.
- Bunge, M. (1979), *Treatise on Basic Philosophy*, Vol. 4, Dordrecht, Reidel.
- Chandler, J. L. R. (1997), Emergence of complex organized natural systems, in *Proceedings ISAS '97*, NIST Special Publication 918, 211-216.
- Changeux, J.-P. (1983), *L'homme neuronal*, Paris, Fayard.
- Cordier, J.-M & Porter, T. (1989), *Shape Theory*, New York, Wiley.
- Damasio, A. (1999), *The feeling of what happens*, New York, Harcourt Brace & C°. Dennett, D. (1990), *La strategie de l'interprete*, Paris, NRF Gallimard.
- Eccles, J.C. (1986), Do mental events cause neural events? *Proc. Royal Soc. Lond.* B227, 411-428.
- Edelman, G.M. (1989), *The remembered Present*, New York, Basic Books.
- Ehresmann(-Bastiani), A. & Ehresmann, C. (1972), Categories of sketched structures, *Cahiers Top. et Geom. Diff.* XIII-2. (Reprinted in *Charles Ehresmann, Oeuvres completes et commentees*, Partie IV-1, Amiens 1984.)
- Ehresmann, A. C. and Vanbremeersch, J.-P. (1987), Hierarchical evolutive systems, *Bull. Math. Biol.* 49, 13-50.
- Ehresmann, A. C. and Vanbremeersch, J.-P. (1991), Un modele pour des systemes evolutifs avec memoire..., *Revue Intern. de Systemique* 5 (1), 5-25.
- Ehresmann, A. C. and Vanbremeersch, J.-P. (1996), Multiplicity Principle and emergence in MES, *SAMS* 26, 81-117.
- Ehresmann, A. C. and Vanbremeersch, J.-P. (1999), *Site Internet'*.
<http://perso.wanadoo.fr/vbm-ehr>.
- Eilenberg, S. & Mac Lane, S. (1945), General theory of natural equivalences, *Trans. Am. Math. Soc.* 58, 231-294.
- Gazzaniga, M. S. (1985), *The social brain*, New York, Basic Books.
- Goffman, E. (1973), *La mise en scene de la vie quotidienne*, Editions de Minuit, Paris. Gray, J. W. (1989), The category of sketches as a model for algebraic semantics, in *Categories in computer science and logic* (Ed. Gray, J.W. & Scedrov, A.), Providence R.I., Am. Math. Soc.
- Hebb, D. O. (1949), *The organization of behaviour*, New York, Wiley.

- Hubei, D. H. and Wiesel, T.N. (1962), Receptive fields..., *J. Physio.* 160 (1).
- Josephson, B. (1998), Extendibility of activities and the design of the nervous system, in *Proceedings third international conference on emergence ECHO III* (Ed. Farre), Helsinki.
- Kan, D. M. (1958), Adjoint Functors, *Trans. Am. Math. Soc.* 89, 294-329.
- Kuhn, T. S. (1972), *La structure des revolutions scientifiques*, Paris, Flammarion.
- Mac Lane, S. (1970), *Categories for the working mathematician*, New York, Springer.
- Malsburg C. (von der) & Bienenstock E. (1986), Statistical coding and short-term synaptic plasticity, in *Disordered systems and biological organization*, NATO ASI Series 20, Springer.
- Matsuno, K. (1989), *Protobiology: Physical basis of Biology*, Boca Raton, CRC Press.
- Minsky, M. (1986), *The society of mind*, New York, Simon & Schuster.
- Paton, R. C (1997), Glue, verb and text metaphors in Biology, *Acta Biotheoretica* 45, 1-15.
- Pavlov, I. P. (1960), *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (G.V. Anrep, Ed. and trans.), New York, Dover. (Originally published 1927.)
- Piaget, J. (1964), Le developpement mental de l'enfant, in *Six etudes de psychologie*, Paris, Ed. Denoel-Gonthier. (Originally published in *Juventus Helvetica*, Zurich 1940.)
- Quine, W. (1960), *Word and object*, Cambridge Mass., The M.I.T. Press.
- Rosen, R. (1986), *Theoretical Biology and complexity*, Academic Press.
- Skinner, B.F. (1938), *The behavior of organisms: An experimental analysis*, New York, D. Appleton-Century C°.
- Thom, R. (1988), *Esquisse d'une Semiophysique*, Paris, InterEditions.
- Vallee, R. (1995), *Cognition et Systeme*, Limonest, L'Interdisciplinaire.
- Walters, R. E. C. (1991), *Categories and computer science*, Cambridge University Press.
- Whinten, A. *et al.* (1999), Cultures in chimpanzees, *Nature* 399, 682-685.
- Wittgenstein L. (1953), *Philosophical Investigations*, Oxford, Blackwell.