Progress in Biophysics and Molecular Biology xxx (2015) 1–13

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Conciliating neuroscience and phenomenology via category theory

Andrée C. Ehresmann [a,*], Jaime Gomez-Ramirez [b]

[a] University of Picardie Jules Verne, France
[b] University of Wisconsin-Madison, United States

A R T I C L E I N F O

Article history:
Available online xxx

Keywords:
Category
Multi-scale self-organized systems
Neuro-cognitive system
Phenomenology

A B S T R A C T

The paper discusses how neural and mental processes correlate for developing cognitive abilities like memory or spatial representation and allowing the emergence of higher cognitive processes up to embodied cognition, consciousness and creativity. It is done via the presentation of MENS (for *Memory Evolutive Neural System*), a mathematical methodology, based on category theory, which encompasses the neural and mental systems and analyzes their dynamics in the process of 'becoming'. Using the categorical notion of a colimit, it describes the generation of mental objects through the iterative binding of distributed synchronous assemblies of neurons, and presents a new rationale of spatial representation in the hippocampus (Gómez-Ramirez and Sanz, 2011). An important result is that the degeneracy of the neural code (Edelman, 1989) is the property allowing for the formation of mental objects and cognitive processes of increasing complexity order, with multiple neuronal realizabilities; it is essential "to explain certain empirical phenomena like productivity and systematicity of thought and thinking (Aydede 2010)". Rather than restricting the discourse to linguistics or philosophy of mind, the formal methods used in MENS lead to precise notions of Compositionality, Productivity and Systematicity, which overcome the dichotomic debate of classicism vs. connectionism and their multiple facets. It also allows developing the naturalized phenomenology approach asked for by Varela (1996) which "seeks articulations by mutual constraints between phenomena present in experience and the correlative field of phenomena established by the cognitive sciences", while avoiding their pitfalls.

1. Introduction

Despite the huge progresses in brain research in the last 25 years, the brain's large-scale organizational principles allowing for the emergence of cognitive abilities like perception, memory, or spatial representation are far from clear. One question we must address to make real progress in the brain/mind problem is: Can we hope to find common processes at the basis of cognition, leading to a new cognitive neuroscience comparable in terms of parsimony and explanatory power with for example, physics?

Mathematical models of brain dynamics have been developed, most often based on non-linear differential equations (Freeman and Vitiello, 2006), dynamical systems theory (Izhikevich, 2006), complex network theory (Bullmore and Sporns, 2009), stochastic variational methods (Friston, 2010) or information theory (Barlow, 1972). They tend to concern particular processes and cannot simultaneously cover the micro, meso and macro levels.

However, despite the diverse nature of cognitive abilities like memory, spatial representation or higher cognitive processes up to consciousness and creativity, they all share the following common properties.

(i) *Synaptic plasticity* (Hebb, 1949): a mental object activates a *neuronal assembly* which operates synchronously and becomes reinforced by Hebb synaptic rule.
(ii) *Degeneracy of the neural code* (Edelman, 1989) a mental object can activate different neuronal assemblies.
(iii) *Structural Core* consisting of a spatially and topologically central sub-graph of the graph of neurons and synapses between them, with many strongly connected hubs (Hagmann et al., 2008). The *Structural Core* plays "a central role in integrating information across functionally segregated brain regions" and "is linked to self-referential processing and consciousness".

* Corresponding author.
E-mail addresses: ehres@u-picardie.fr (A.C. Ehresmann), jd.gomezramirez@gmail.com (J. Gomez-Ramirez).

http://dx.doi.org/10.1016/j.pbiomolbio.2015.07.004
0079-6107/© 2015 Published by Elsevier Ltd.

(iv) *Multi-temporality modular self-organization* of the neural system, with modules of different sizes, working at different rhythms.

Using categorical tools, these properties allow constructing a 'dynamic model' MENS (for *Memory Evolutive Neural System*) of a neuro-cognitive-mental system of which we give an outline in this article. MENS proposes a common frame to study neuronal and mental processes up to the development of higher order cognitive processes, at different levels of description and across different timescales, with their temporal becoming, "to acknowledge the openness of this becoming" (Kauffman and Gare, 2015).

MENS was introduced by Ehresmann and Vanbremeersch (2001, 2007) to account for neuroscientists' results, for instance Edelman's work in degeneracy; and it has also benefited from phenomenology, in particular the works of Husserl, Brentano and Merleau-Ponty. Crucially, some of the mathematical notions brought by MENS were later 'found' to have a neurological correlate. For instance: (i) cat-neurons (introduced in the nineties) have exactly the properties that the neuroscientist Buzsaki gave (in 2010) to his "reader neurons" and their hierarchy could describe his "neuronal syntax"; (ii) the Archetypal Core was introduced in (Ehresmann and Vanbremeersch, 2001) for studying consciousness, and it is only in 2008 that its neuronal base (the structural core mentioned above) has been discovered.

In MENS, the mental "supervenes" on the neural (through iterated complexifications), so that it relates to the neurophenomenology introduced by Varela (1996) while avoiding the pitfalls indicated by Bayne (2004). In particular, the degeneracy property avoids "isomorphism between neural and mental" because it implies that a mental object has multiple neuronal realizations.

The paper is structured as follows: Section 2 recalls the basic notions of category theory and how they provide an approach to the notion of mathematical structure and to universal properties. Section 3 shows how the notions of direct and inverse limits (Kan, 1958), of (hierarchical) evolutive systems and the complexification process (Ehresmann and Vanbremeersch, 1987) lead to the development of a methodology for studying the evolution of multi-scale self-organized systems. MENS is constructed in Section 4 by iterated complexifications of the (evolutive system NEUR modeling) the neural system; its local and global dynamics in their temporal becoming are studied in Section 5. In Section 6, we stress the role of the Archetypal Core at the root of both the emergence of higher cognitive processes and of phenomenological experiences. Section 7 analyzes how the notions introduced deal with empirical phenomena like compositionality, productivity and systematicity of thought (Aydede, 2010) and how the MENS methodology relates to philosophical problems such as the well-known confrontation connectionism vs. classicism, or the development of a neurophnomenology.

2. Category theory

Category theory is a domain of mathematics, introduced by Eilenberg and Mac Lane in the forties (Eilenberg and MacLane, 1945) to relate algebraic and topological constructs. Category theory has a unique status, at the border between mathematics, logic, and meta-mathematics. Crucially, it resorts to relational mathematics, since what is important in a category is not the "structure" of its objects per se, but the relations between them. In the late fifties, its foundational role in mathematics was made apparent, in particular through the introduction of adjoint functors and (co) limits by Kan (Kan, 1958), the theory of species of structures and of local structures by Charles Ehresmann (Ehresmann, 1958), and the notion of abelian categories as a basis for homology (Grothendieck, 1957). Later its role in logic was emphasized by several authors: for example, in the theory of *topos* developed by Lawvere and Tierney (Awodey et al., 2009), and in the *sketch theory* developed by Ehresmann (Borceux, 2009). It makes a general concept of structure possible, and indeed it has been described as mathematical structuralism, providing a single setting unifying many domains of mathematics.

Category theory tries to uncover and classify the main operations of the "working mathematician"; for instance defining a general notion of sub-structure, of quotient structure, of product,… valid as well for sets, groups, rings, topological spaces,… Mathematical activity, here, reflects some of the main operations that humans do for making sense of the world: distinguishing objects (a tree, a fruit,…); formation, dissolution, comparison, and combination of relations between objects (the fruit is linked to the tree, these fruits have the same color, one fruit is larger than another,…); synthesis of complex objects from more elementary ones (binding process) leading to the formation of hierarchies (complexification process); optimization processes (universal problems); classification of objects into invariance classes (formation of concepts). As all these operations are at the root of our mental life, and also of science, it quite naturally follows that category theory can be successfully applied to different scientific domains (Spivak, 2014), in particular computer science, in the foundations of physics for studying quantum field theories, and in biology, see for example the seminal work of Robert Rosen (Rosen 1958) and recent contributions in the field of theoretical biology (Letelier et al., 2006), (Gatherer and Galpin, 2013).

2.1. Graphs. The graph of neurons

Graphs have been used to represent networks of any nature: cellular networks, social networks, the internet… Here we define a *graph* G as a set G_0 of objects A, B,…, called its *vertices*, and a set of oriented edges (or *arrows*) between them; an edge *f* from A to B is represented by an arrow *f*: A → B. It is possible to have several arrows with the same source A and the same target B, and even 'closed' arrows (the source and target are identical). Let us remark that the term 'graph' is often restricted to the case where there is at most one arrow from a vertex to another, in which case the graph can be represented by a binary matrix.

A *path* of the graph from A to B is a sequence of consecutive arrows

$$(f_1, f_2, \ldots, f_n) \text{ with } f_1 : A \to A_1 \cdot f_2 : A_1 \to A_2, \ldots, f_n : A_{n-1} \to B.$$

The paths of G form the *graph of paths of* G, denoted P(G): it has the same vertices as G but its arrows from A to B are the paths of G from A to B. We identify G with a sub-graph of P(G) by identifying an arrow *f* to the path (*f*) with *f* as its unique arrow.

If G and G′ are two graphs, a *homomorphism p* from G to G′ associates to each vertex A of G a vertex *p*(A) of G′, and to each arrow *f* from A to B an arrow *p*(*f*) from *p*(A) to *p*(B).

Example: *The neuronal graph* at an instant *t*: A vertex $N_t = (N, n(t))$ models the state at *t* of a neuron N with its activity *n*(*t*) at *t* (measured by its instantaneous firing rate). An arrow $f_t = (f, p(t), s(t))$ from N_t to N'_t models a synapse *f* from N to N′, labeled by its *propagation delay* *p*(*t*) around *t* and by its *strength* *s*(*t*) to transmit an action potential from N to N′. The strength (negative if the synapse is inhibitory) varies according to *Hebb rule*: it increases if the activations of N and N′ are correlated. The graph of paths of the neuronal graph will be at the root of our model; the *propagation delay* of a path is defined as the sum of those of its factors; and its *strength* as the product of their strengths.

Graphs and their paths are not sufficient to account for the fact that several paths from N to N′ may play equivalent operational roles. For instance, different synaptic paths may transmit the same activation from neuron N to neuron N′. The notion of a category enriches that of a graph by allowing a comparison of paths to distinguish ‘operationally equivalent paths’.

2.2. Categories and functors

A *category* is a graph equipped with an internal composition associating to a 2-path (f, g) where f: A → B and g: B → C the arrow fg: A → C *composite* of the arrows f and g. The vertices are also called objects of the category and the arrows can also be called *morphisms*, or more simply *links.*

The composition satisfies 2 conditions:

(i) it is associative: for a path (f, g, h) we have $f(gh) = (fg)h$. It follows that each path, say $(f_1, f_2,\ldots, f_n)$ (whatever its length) has a unique composite (whatever its 2–2 decomposition) denoted $f_1 f_2 \ldots f_n$.
(ii) for each object A there is an arrow id_A from A to A, the *identity* of A, whose composite with any arrow of source or target A is identical to f.

A *functor p* from a category C to a category C′ is a homomorphism of graphs which respects the composition and the identities, so that

$$p(fg) = p(f)p(g) \quad \text{and} \quad p(\mathrm{id}_A) = \mathrm{id}_{p(A)}.$$

2.3. Examples of categories

(i) A *monoid* is a category with a unique object.
(ii) To a *poset* (E, <) we associate a category admitting E as the set of its objects and with a morphism from A to B if and only if A < B for the order; the categories so associated to posets are characterized by the fact that there is at most one morphism between 2 objects.
(iii) A *groupoid* is a category in which each morphism f: A → B has an inverse f': B → A such that the composites ff' are $f'f$ are identities; in particular a group is a groupoid with a unique object.
(iv) To each graph G, we associate the *category of paths* of G: it is obtained by equipping the graph of paths of G with the convolution of paths; the identity of an object A is the ‘void’ path from A to A. Each category is a quotient of the category of paths of its underlying graph by the equivalence relation on paths: “two paths are equivalent if they have the same composite”.
(v) Given a graph G and a category C we define the category C^G whose objects are homomorphisms P from G to the underlying graph of C, and the morphisms from P to another homomorphism P′ are the *natural transformations u*: P → P′ where u is a map from the set of objects of G to C such that $P(x)u(j) = u(i)P'(x)$, for each $x: i \to j$ of G..

2.4. Categories of structured sets

Category theory helps giving a precise definition of structure. For Mac Lane (MacLane, 2010), “a structure is essentially a list of operations and relations and their required properties, commonly given as axioms, and often so formulated as to be properties shared by a number of possibly quite different specific mathematical objects”. For Landry (Landry, 1999), “category theory, in virtue of its ability to organize our talk about both structures and the structure of structures, ought to be taken as a framework for mathematical structuralism”. Here we consider how to give a general definition of structures on sets, by following the ‘categorical’ translation that Charles Ehresmann (Ehresmann 1957) has given of Bourbaki's theory of set-structured systems.

We denote by *Set* the category whose objects are sets (e.g. of a universe to avoid size problems which we do not raise here), the morphisms from A to B being the maps from A to B, with their usual composition. Mathematical structures over sets give rise to categories with a ‘forgetful’ functor toward *Set*, for instance: the category *Group* whose objects are groups and the morphisms are homomorphisms between groups, the functor associating to a group its underlying set (thus ‘forgetting’ the group structure); the category *Top* whose objects are topological spaces and the morphisms are continuous maps between them; the category *Cat* whose objects are (small) categories and the morphisms are the functors between them,… Such categories of structured sets have been characterized by Ehresmann, under the name homomorphism categories as follows:

Definition. A *homomorphism category* over *Set* is a category *C* equipped with a functor *p* from *C* to *Set* satisfying the 2 conditions:

(i) p is faithful: if c and c' are morphisms of *C* from A to B and if $p(c) = p(c')$, then $c = c'$.
(ii) Transport by isomorphism: If A is an object of *C* and f a bijection from p(A) to E, there exists an isomorphism c: A → A′ in C such that $p(c) = f$; then A′ is called the structure transported from A by f.

2.5. Universal properties

To say that A has a *universal property* means that A is a most efficient (or “universal”) solution to a specific problem. This notion is made precise in the categorical setting, in relation with the notion of adjoint functors (Kan, 1958). Universal constructions are ubiquitous both in Mathematics and in its applications; for instance, (Phillips and Wilson, 2010) write: “All systematic and quasi-systematic properties of human cognition are just instances of universal constructions”.

By Yoneda Lemma, an object A of a category is characterized by the morphisms to (or out of) A. Thus, a universal property of A will be recognized through properties of such morphisms, without necessitating an explicit construction of A, and it will determine A up to an isomorphism. For instance, the characterization of products is the same in any category, be it the category of sets, of groups, of modules, of topological spaces, and so on.

The generic example of a universal property for an object of C is given by an *initial object* A of C: it is an object with exactly one morphism from A to any other object of C.

By duality (through reversing the arrows) we obtain a *final object* to which arrives exactly one morphism from any other object. Each other universal property can be converted into an initial or a final object of an appropriate category. In the next section, we explain how it is done in the case of colimits and limits (direct and inverse limits in the sense of Kan) and this rationale will be extensively used in the following sections.

3. How to model a multi-scale evolutionary system?

Categorical models of concrete systems generally are concerned with the invariant structure of the system; an instance is the Metabolic-Repair (M-R) systems introduced by Rosen (1958). An M-R system can be thus represented by a simple category.

Evolutionary multi-scale systems have multiple components of different complexity levels which, moreover, can vary over time, with loss, addition, combination or decomposition of components (the "standard changes" in the terms of Thom, (Thom, 1988)); the neuro-cognitive-mental system studied later is an example. Such systems raise 2 problems: defining a hierarchy of components, and introducing time as an operative tool. The first problem leads to the introduction of hierarchical categories, and the second to the notion of an Evolutive System consisting of a family of categories indexed by time. Both unite into the notion of a Hierarchical Evolutive System.

3.1. Colimits and limits

A particular kind of universal property leads to the notion of limits and colimits, which provide general mechanisms for combining structures and answer the problem: how do simple objects bind together to form a complex object with emerging properties (for instance a wall acquires a property of enclosing that the heap of bricks of which it is formed has not) and how such complex objects interact? In MENS, colimits will be used to explain how the dynamics of the neuronal system produces mental representations.

More abstractly, let C be a category. A *pattern* in C is a homomorphism P from a graph sP into C (if sP is a category, it is called a *diagram* in C). A *cone* with basis P and vertex N is a natural transformation from P to the homomorphism from sP to C constant on N. These cones are the objects of a full sub-category of C^P. A *colimit of* P is an initial object of this category of cones.

Let us develop these definitions, using the more concrete terminology of (Ehresmann and Vanbremeersch, 2007).

Definition.

(i) A *pattern* P in C consists of a family $(P_i)_i$ of objects P_i of C and some distinguished morphisms between them. A *cone* with basis P and vertex N, also called a *collective link* from P to N, consists of a family $(s_i)_i$ of links s_i from P_i to N, which satisfies the following equations:

$s_i = fs_j$ for each distinguished link f from P_i to P_j.

(ii) An object M is a colimit of P if there is a collective link $(c_i)_i$ from P to M satisfying the universal property: Any collective link $(s_i)_i$ from P to N factors through a unique link s from M to N such that $s_i = c_i s$ for each i. We say that s binds (s_i).

A pattern P may have no colimit; if it has a colimit it is unique up to an isomorphism. The fact that P admits M for colimit has for consequence the *systematic* association, to each collective link from P to N, of a morphism from M to N binding it. Thus, *the construction of colimits implicates a systematicity property* to be used later.

An (*inverse*) *limit* of P in a category C is defined as a colimit in the category C^{op} opposite to C obtained by inversing the direction of the arrows.

Examples. 1. Let P be a pattern without distinguished links; its colimit is called the coproduct of the family $(P_i)_i$ and its limit the product of (P_i). In Sets the coproduct is the disjoint union of the P_i's, and the limit is $\{(p_i)_i | p_i \in P_i \text{ for each } i\}$.

2. In the category defining an order, the colimit of a pattern P is also the coproduct of the family $(P_i)_i$ as well as the lower upper bound of this family; its limit is the greatest lower bound of the family. For instance, in the category defining the order on the parts of a set E, the colimit is the union of sets and the limit their intersection.
3. If P consists of 2 morphisms (f: A → B, f': A →°B′) with the same source A, its colimit is called their *pushout*; in *Sets* the *pushout* is the quotient of the disjoint union of B and B′ by the equivalence generated by $f(a)$°~°$f'(a)$ for each a in A. The limit of 2 morphisms with the same target is called their *pullback*.

In the sequel, the *colimits* will make possible to represent the hierarchy of complexity levels of a multi-scale system.

3.2. The primary role of time: evolutive systems

Generally speaking, when categories are used in relation with biological systems, the idea is to have a category modeling the invariant structure of the system (for an in depth study of this idea, see (Rosen, 1985a)). In MENS the perspective is different for we are interested in the evolution of the system with its dynamic, studied in its 'becoming' rather than its 'being'. Importantly, time intervenes under different forms: continuous 'clock-time' used to describe the dynamic of the system and to quantify the propagation and activation delays; the measure of change (as in Augustine of Hippo); discrete timescale specific to each co-regulator for delimiting its successive steps.

To account for this, the system (be it the neural system or the neuro-cognitive-mental system) will not be modeled by a unique category, but by an Evolutive System.

An Evolutive System **K** consists of:

(i) A family of categories indexed by the timeline T of the system (an interval of the positive reals line); each of these 'configuration' categories K_t represents the state of the system at a given time t of its existence.
(ii) For each time $t' > t$, a functor $k_{tt'}$ from a subcategory $K_t°$ of K_t to $K_{t'}$ called transition; $K_t°$ models the elements 'still persisting' at t', and the functor their change from t to t'. These transitions satisfy a transitivity condition (cf. Appendix).

We define a *component* C of the Evolutive System **K** as a maximal family $(C_t)_t$ of objects C_t of the categories K_t related by the transition functors (cf. Fig. 2); C_t is called the state of the component C at t, and C is the (dynamic) trajectory of its different states. A *link* between components is defined similarly by the successive states of a morphism between successive states of the components.

A configuration gives a snapshot of the system in its *simultaneity* at a given date t; the transitions account for the *succession* of such snapshots and measure the changes between them (but not how the change has been produced). On the other hand, components and links between them keep track of their successive states, thus giving a '*transversal*' view allowing for presenting the dynamic. Components and links defined on an interval J of T form a category

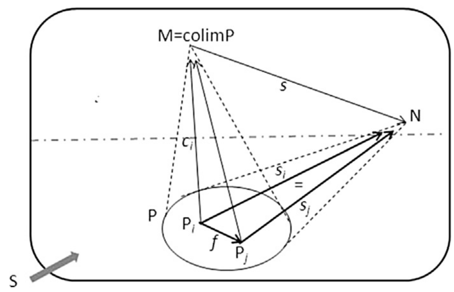

Fig. 1. Collective link and colimit.

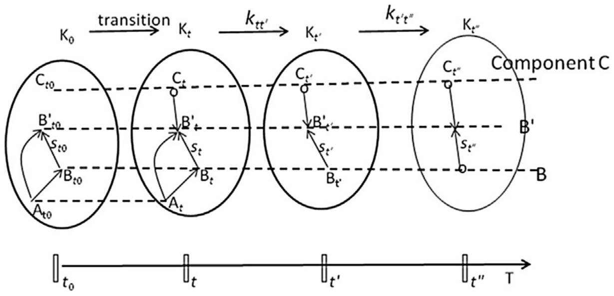

Fig. 2. An Evolutive System **K** and some of its components.

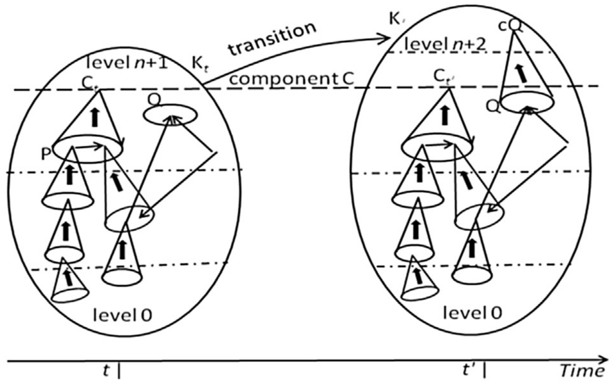

Fig. 3. A Hierarchical Evolutive System, with a ramification of C. The transition corresponds to a complexification which binds Q into cQ.

K_J; the categories K_J form a sheaf of categories over T. We often work in such categories K_J in particular when we speak of a pattern of components of **K** and its colimit (if it exists).

Example. The neural system will be modeled by an *Evolutive System* NEUR, called the *Neural Evolutive System* and denoted by NEUR. At each time t of the life of the system, the configuration category $NEUR_t$, called the *neuronal category at t.* is the category of paths of the neuronal graph at t (defined in Section 2.1). Its objects N_t model the neurons N existing at t with their activity $n(t)$, the morphisms (also called *links*) the synaptic paths between them, with their propagation delay and strength at t. This neuronal category varies over time: the transition from t to a later time t' is the functor from a sub-category of $NEUR_t$ to $NEUR_{t'}$ which associates to N_t its new state at t' provided that N still exists at t', and similarly for the links. The components of NEUR, called 0-*cat-neurons* model the neurons N with their activity n during their life, and the links between them model synaptic paths with the variation of their propagation delays and strengths. Thus, NEUR gives a faithful dynamic model of the neural system, reflecting both the 'birth' and 'death' of neurons or synapses and their dynamic during their life. But it does not explain how the dynamic is internally generated (for this cf. Section 5).

The dynamic of an evolutionary system such as the neural system results from different 'physical' operations (e.g., formation of a new synapse), which necessitate the activation of some links of the system and require a specific 'duration'. To reflect this property of links in an Evolutive System **K** we give a functor d_t from K_t to $\mathbf{R}_+$ which associates to a morphism f_t of K_t a real number $d_t(f_t)$ called its *propagation delay* at t; moreover the morphisms of K_t are divided into 2 classes: those which are said to be *active at t* and those which are not. In NEUR a link is active at t if it models a synaptic path from N to N′ transmitting an action potential from N to N′ around t.

3.3. Hierarchical evolutive systems. Complexification process

To account for the different complexity levels of its components, an evolutionary multi-scale system, such as the neuro-cognitive-mental system MENS, will be represented by a *Hierarchical Evolutive System*: that is an Evolutive System in which the configuration categories are hierarchical, and the transitions respect the level.

Definition. A category is *hierarchical* if its objects are partitioned into a finite number of levels numbered 0, 1,…, m such that each object C of level $n+1$ is the colimit of at least one pattern (of interacting objects) contained in the levels $\leq n$. Then C has also a *ramification* down to level 0 (cf. Figure 3). The *complexity order* of C is the length of the shortest ramification of C; it is thus less or equal to the level of C. A *Hierarchical Evolutive System* is an Evolutive System in which the configuration categories are hierarchical and the transitions respect the level.

In a Hierarchical Evolutive System, the transitions are composites of elementary transitions for which the change from t to t' results from operations of the following types: loss or decomposition of some complex objects, binding of some patterns Q leading to the formation, or preservation if it exists, of a colimit cQ of Q.

In categorical terms, this operation can be modeled by the **Complexification Process.** Given a procedure Pr on the category K_t the problem of constructing a category in which the objectives of Pr are realized has a universal solution, called the complexification of K_t with respect to Pr. (For a formal presentation of the complexification process, where a procedure is modelled by a pro-sketch, cf. Appendix.)

The transition from t to t' is obtained by complexification if there is a procedure Pr on the configuration K_t such that the configuration at t' be the complexification $K_{t'}$ of K_t with respect to Pr. The complexification can be explicitly constructed (Ehresmann and Vanbremeersch 1987); in particular its objects are the new states of the components not suppressed by Pr, as well as new components cQ which become the colimit of patterns Q specified by Pr (if 2 patterns Q and Q′ have the same operational role, we take $cQ = cQ'$). Let us note that, if the HES models a 'physical' (or biological) system, the realization of the 'physical' changes induced by the complexification process requires some duration (from t to t' above). This point will be conveniently explained, in the case of MENS, in Section 4.

The complexification procedure may also ask for the formation of limits of some patterns, then we speak of a *mixed complexification.* Its construction is more complicated (for more details, see (Ehresmann and Vanbremeersch 2007)).

3.4. Multiplicity principle MP at the base of non-reductionism and emergence

MP is a kind of 'flexible redundancy' which formalizes the degeneracy of the neural code (Section 1) and extends it in the frame of a hierarchical evolutive system.

Definition. In a hierarchical category, an object M of level $n+1$ is *n-multi-faceted* if it is the colimit of several non-isomorphic patterns of levels $\leq n$ such that there is no cluster of links (cf. Appendix) between them binding into the identity of M. If the category admits such multi-faceted objects, we say that it satisfies the *Multiplicity Principle*.

In a HES, a component M is multi-faceted if it has multi-faceted states. Then with time M takes its own complex identity (or 'individuation') with possibility of addition or loss of decompositions (hence also of ramifications) and 'switches' between them (cf.

Please cite this article in press as: Ehresmann, A.C., Gomez-Ramirez, J., Conciliating neuroscience and phenomenology via category theory, Progress in Biophysics and Molecular Biology (2015), http://dx.doi.org/10.1016/j.pbiomolbio.2015.07.004

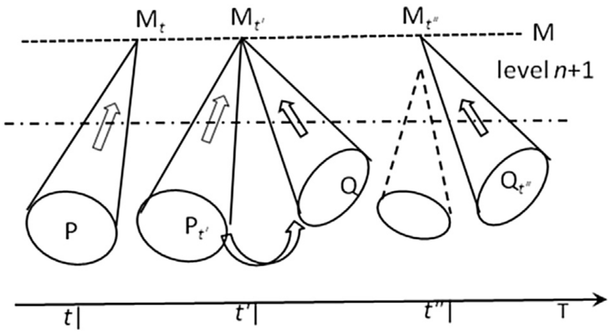

Fig. 4. The component M takes its individuation over time.

Figure 4). This gives flexibility to the system, in particular allowing for the development of a Memory adapting to changes.

MP is also at the root of emergence properties. Indeed, among the links from M to M′ there are *n-simple links* obtained as the binding of a cluster of links between patterns P and P′ of levels $\leq n$ having M and M′ as their colimits. MP allows for the emergence of *n-complex links* obtained as composites of *n*-simple links binding non-adjacent clusters (cf. Fig. 5). Such links represent non-local properties 'emerging' at the level $n + 1$ from the global structure of the lower levels. (For more details, cf. Appendix and Section 4.3.)

The existence of complex links is essential in the proof of the following theorems (Ehresmann and Vanbremeersch, 1996, 2007).

Complexity Theorem. In a HES, MP is necessary for the existence of components of complexity order >1; without it, all components are colimit of a pattern of level 0.

Emergence Theorem. MP is preserved by a complexification process. It is necessary for the emergence, through iterated complexifications, of components of increasing complexity orders and of complex links between them.

Iterated Complexification Theorem. With MP, two successive complexifications may not be reducible to a unique one.

The proof of this last theorem depends on the emergence, in the first complexification, of complex links which introduce "change in the conditions of change" (Popper, 1972). As proved in (Ehresmann and Vanbremeersch, 2007) it has for consequences the mixing of Aristotelian formal, material and efficient causes which, for Rosen (1985b), distinguishes organisms from pure mechanisms.

4. Construction of MENS: how the mental emerges from the brain

The binding process is ubiquitous in evolution and it raises the following problem: how can simple objects bind together to form "a whole that is greater than the sum of its parts"? what are the different patterns of "parts" leading to the same whole, and what are the simple and complex interactions between such complex wholes? This problem arises in Neuroscience, where it has been emphasized by (Malsburg and Bienenstock, 1986), to explain how mental objects can emerge from exchanges in the brain, and how they interact. Our aim is to show how the notions of colimits and complexification help solving it.

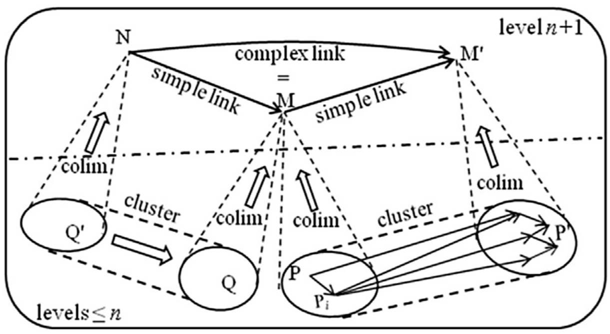

Fig. 5. Cluster, simple and complex links.

4.1. Properties of the neural system

As recalled in the Introduction, a mental object (e.g. the mental image of a stimulus) activates a synchronous assembly of neurons; but how do such assemblies interact? The mental object could be represented by a neuron if there was a neuron N 'binding' the assembly, in the sense that N and the assembly as such have the same activating role on other neurons. For instance (Hubel and Wiesel, 1977) have shown the existence of neurons representing a segment or an angle; and there are also neurons representing more complex but very familiar objects. In (Gomez and Sanz, 2009, Gómez-Ramirez and Sanz, 2011, 2013, Gomez-Ramirez, 2014) it is shown that hippocampal place cells can be modeled as the colimit of a pattern of grid cells. It ought to be remarked, however, that "grandmother neurons" (Barlow, 1972), binding the neuronal assembly activated by a mental object, cannot be expected to be found. Moreover, the *degeneracy* property of the neural code emphasized by Edelman (Edelman, 1989) means that, depending on the context, the same mental object can activate structurally different assemblies.

How to represent the mental object as such and to determine how it interacts with neurons and other mental objects? To answer these questions, we model the neuro-cognitive-mental system by the *Memory Evolutive Neural System* MENS (Ehresmann and Vanbremeersch, 2007), whose higher levels emerge as a superstructure over the 0-level infrastructure NEUR, more precisely, MENS is a HES generated by the Neural Evolutive System NEUR (cf. Section 2) through successive complexification processes. Its components are dynamic objects, called *category-neurons* (abbreviated in *cat-neurons*) which represent the temporal trajectories of more and more complex mental objects and processes. The idea is that a cat-neuron of level 1 'binds', or in MENS nomenclature, becomes the colimit of, each pattern of NEUR modeling a synchronous assembly of neurons activated by the mental object. The construction of the complexification will indicate the 'good' (simple and complex) links between them. Thus, we can speak of patterns of cat-neurons of level 1 and iterate the construction. Successive iterations lead to higher level cat-neurons representing more and more complex mental objects.

4.2. Formation of (multi-faceted) cat-neurons as colimits

An assembly of neurons A is modeled by a pattern P in the Neural Evolutive System NEUR. The pattern consists of a family of 0-cat-neurons P_i interconnected by some distinguished links *f* from P_i to P_j (representing the synapses through which the neurons of A transmit their activation to each other). If the assembly is *synchronously activated* at *t* (meaning the activities of all its neurons increase), the same is valid for P since the passage to NEUR preserves the activities.

If a stimulus S activates A, this assembly can synchronously activate a neuron N at a time *t* if there are links from its different neurons to N which all transmit the activation of P_i to N at the same time *t*, taking account of the interactions of the neurons in A. By construction of NEUR, this 'activation' operation is modeled in NEUR by the data of a collective link $(s_i)_i$ from the pattern P to N (cf. Section 3). The set of equations:

$s_i = fs_j$ for each distinguished link f from P_i to P_j

imply that the propagation delay of s_i is the sum of those of f and of s_j. If there is such a collective link, we say that P acts as a *polychromous* pattern in the sense of (Izhikevich et al., 2004).

If the stimulus S is repeated or persists, the distinguished links of A, being simultaneously activated by S, are strengthened via Hebb rule (Hebb, 1949). Transferred in NEUR it means that P takes an identity of its own to act as a polychronous pattern. The long-term memory of S will be recorded in MENS by a cat-neuron M 'binding' P, which becomes the colimit of P (cf. Fig. 1). For instance S could be a rectangle and P the pattern consisting of (the 0-cat-neurons activated by) its sides and vertices, with distinguished links from a vertex to the sides containing it.

The degeneracy property of the neural code asserts that the stimulus S can activate more or less different polychronous assemblies of neurons, simultaneously or at different times; though these assemblies can be structurally different and not even well connected, they all have the same operational role since they correspond to the same mental object. In MENS, the cat-neuron M recording S must represent the invariant common to all these patterns; hence it must bind each one of them; thus it is *multi-faceted,* and MENS satisfies the *Multiplicity Principle*.

Initially constructed to bind a particular pattern P, the cat-neuron M later takes its own individuation as a component of MENS and can even disassociate from P at a later time t' (cf. Fig. 4). It is not a rigid memory (as in a computer), but a flexible dynamic representation which adapts to changing situations, so that S can later be recognized through the activation of M under different forms, even new forms not yet met (as long as the change is progressive enough).

4.3. The hierarchy of cat-neurons and their links

The compositionality of mental representations means that complex mental objects are constructed by binding together patterns of simpler ones. To ensure it in MENS, we construct a hierarchy of cat-neurons and links between them, from the level 0 of 0-cat-neurons (that is, NEUR) up, by successive complexification processes (Section 3). More explicitly, once constructed the cat-neurons of level less or equal to $n > 0$ and their links, the cat-neurons of level $n+1$ are obtained as colimits of polychronous patterns of cat-neurons of levels less or equal to n, and there are two kinds of links between them (cf. Fig. 5):

(i) *Simple links.* Let M and M′ be 2 cat-neurons of level $n+1$, binding respectively patterns P and P′ of levels $\leq n$. A (P, P′)-*simple link* from M to M′, or *n-simple link,* is a link binding a cluster (cf. Appendix) G of links from P to P′. Such a link is *systematically* associated to G, using the universal property of colimits. Being deducible from interactions between components of P and P′, it just translates at the level $n+1$ of cat-neurons the fact that P can coherently activate P′ through G.

(ii) *Complex links.* A composite of n-simple links binding adjacent clusters is n-simple. However, because of the existence of multi-faceted cat-neurons M, there are also *n-complex links* which are the composites in MENS of n-simple links binding non-adjacent clusters separated by a switch, for instance (cf. Fig. 5) an n-complex link from N to M′ composite of a (Q′, Q)-simple link with a (P, P′)-simple link, where P and Q are structurally different non-connected (cf. Appendix) decompositions of M. They are not discernible 'locally' through the lower components of N and M′, but *emerge* at the cat-neuron level $n+1$ as an outcome of 'global' properties of the lower levels. These complex links reflect global properties of the lower levels which are not observable locally at these lower levels. They are at the root of the emergence of mental representations of increasing complexity.

By construction, the cat-neurons are partitioned into different levels, so that each cat-neuron M of level $n+1$ binds at least one pattern P with values in the full sub-category whose objects are of level $\leq n$; and some M are multi-faceted. Thus, the categories $MENS_t$ are *hierarchical* and satisfy the *Multiplicity Principle*, and MENS is a Hierarchical Evolutive System satisfying MP.

Using the universal property of the complexification, it has been shown (Ehresmann and Vanbremeersch, 2009) that the propagation delays and strengths of the links in NEUR extend to the links in MENS, and there is an *extended Hebb rule*: The strength of a link from M to M′ increases if the activities of M and M′ vary in the same direction.

4.4. Complexity order of a cat-neuron modeling a mental object

Let us emphasize that a cat-neuron M of level $n+1$ is not a formal symbol but a dynamic adaptive multi-faceted representation of a mental object or process which, once formed at a time t, takes its own individuation up to its 'death'. Indeed, by definition of the hierarchy, M admits at least one *ramification* down to the level 0, obtained by taking a decomposition P into a polychronous pattern of cat-neurons of levels $<n+1$, then a decomposition of each component P_i of P into a polychronous pattern of cat-neurons of strictly lower levels and so on, down to patterns of 0-cat-neurons, which form the *base of the ramification*. By construction of NEUR, the patterns in this base model assemblies of neurons which form what we call the *neuronal base* of the ramification.

Because of the multiplicity principle, the cat-neuron M maybe multifaceted and, over time, acquire or loss some ramifications. The number of structurally different ramifications of M measures its *entropy* or amount of variability. The ramifications of M have not all the same length. For instance (the cat-neuron representing) a cube can be directly decomposed into its sides; or first decomposed into its faces, and then each face decomposed into its sides. The *complexity order* of M is the smallest length of a ramification; it is less or equal to the level of M; and it measures the smallest number of steps sufficient for a later activation of M.

By *activation* or recall of (the mental object represented by) M at a time t we mean the unfolding of one of its *ramifications* R down to its base B of level 0 and activation of the patterns of 0-cat-neurons in this base, that corresponds to the *physical activation* of the neuronal base of R. At each step of the construction of a ramification, there is a choice between various (possibly non-connected) decompositions, so that the activation of M has several freedom degrees leading to *multiple physical realizabilities* (in the sense of (Kim, 1998) as hyper-assemblies (i.e. assemblies of assemblies of … assemblies) of neurons. Each operation having some duration (Section 3), the recall of M requires an *activation delay,* which increases with the length of R, hence with the complexity order of M. Activation delays will play a major role in the development of higher cognitive processes. The successive states of M can be activated or not and, when activated, realized under anyone of its ramifications.

Remark. The entropy and the complexity order of a cat-neuron are two different measures of its intricacy, where both denote in different ways to its internal richness and flexibility. In particular, cat-neurons with both a higher complexity order and many ramifications will play an important motor role in the formation of higher cognitive processes.

Since MENS satisfies MP, the theorems of Section 3 can be applied, in particular: In MENS, over time there is emergence of cat-neurons of increasing complexity order, representing more and more complex mental objects or processes. These cat-neurons represent mental states which *supervene* on brain physical states, with multiple physical realizabilities (as explained above), agreeing with Kim's *mental causation.*

5. Local and global dynamics of MENS

5.1. Dynamic of the neural system and its structural core

The neural system has a multi-temporality modular organization, with modules of various types modulating its global dynamics. Successive experiences of any kind, be they sensory, proprioceptive, motor, affective, cognitive, are memorized to be later recalled in analog circumstances. They are processed by specific 'modules' or areas of the brain, for example: small specialized "treatment units" (Crick, 1995) such as the visual centers processing color, to large areas such as the hippocampus, the nuclei of the emotive brain (brain stem and limbic system), or Crick's "conscious units" in the associative cortex.

The dynamic of the neural system depends both on external stimuli and on its own internal activity. For instance: "cortical activity cannot be considered as being driven by the senses, but sensory inputs rather seem to modulate and modify the internal dynamics of cerebral cortex" (Destexhe, 2011). A main integrating role is played by what (Hagmann et al., 2008) call the *Structural Core*:

> "This complex of densely connected regions in posterior medial cortex is both spatially and topologically central within the brain. Its anatomical correspondence with regions of high metabolic activity and with some elements of the human default network suggests that the core may be an important structural basis for shaping large-scale brain dynamics."

5.2. The self-organized dynamic of MENS

The modular self-organization of the neural system is reflected to NEUR and extended to MENS. Thus MENS becomes a *Memory Evolutive System* in the sense of (Ehresmann and Vanbremeersch, 2007), meaning that it is a HES satisfying MP and equipped with a sub-HES called its *Memory* and a network of functional sub-systems called *co-regulators* which operate with the help of the Memory and whose possibly conflicting interactions modulate the dynamic.

(i) The *Memory* is a hierarchical evolutive sub-system MEM of MENS. A cat-neuron in MEM, called *record*, models a mental object associated to an item (external object or signal, past event, internal state, behavior, sensory-motor or cognitive process, habit,…). As any cat-neuron, a record is a dynamic entity which can be 'recalled' through the activation of any of its ramifications. A record takes its own individuation, thus remaining flexible enough to adapt to changing situations; thus it acts as a robust, flexible and adaptive dynamic 'memory' up to its 'death'. MEM develops over time, with emergence of records of increasing complexity (cf. Emergence Theorem) and possible loss of some records. Some records, called *procepts*, model procedures of different kinds (behaviors, automatic processes,…). A procept Pr is associated to a pattern E modeling the effectors through which Pr can be realized, and Pr is the (inverse) limit of this pattern E.

(ii) The *co-regulators* are evolutive subsystems acting as internal regulatory organs. A co-regulator is based on a specific module of the brain, meaning that its cat-neurons have ramifications with their neuronal bases in this module (so that they model distributed hyper-assemblies of neurons of the module). It operates stepwise at its own rhythm, with a specific function (e.g. treating colors) characterized by the kind of links of MENS it may receive or send; in particular it has a differential access to MEM to recall its *admissible procepts* of which it can directly activate the effectors.

5.3. The local dynamic of the co-regulators and their interplay

Each co-regulator operates stepwise as a "hybrid dynamic system", with 2 temporalities: its own discrete timescale which delimits its successive steps (reflecting its 'internal' time), and the shared continuous 'clock-time' which allows measuring the objective length of these steps.

(i) Local dynamic of a co-regulator CR during one step from t to t′.

The partial information accessible to CR during its step consists of the links of MENS which activate some components of CR during the step; they are the components of an Evolutive System **L** with timeline the interval J = [*t*, *t*′[, called the *landscape of* CR on J. This landscape, which acts as a transient working memory for CR during the step, can be thought as the '1st person' perspective of CR during its 'actual present' J. For instance if CR models a brain module treating colors and an object S is presented to the system during J, the landscape will only contain information on the color of S.

Formally, a component *b* of the landscape **L** is a link from a cat-neuron B to a component of CR which remains active during the step; the links in L from *b* to another component *c* of L correspond to commutative squares (*b*, *c*, *f*, *g*) with *f* in CR (cf. Fig. 6).

An admissible procept Pr is selected through a component *pr* of **L** and the effectors of Pr are activated (via the commands *e*). The activation of the effectors through one of their neuronal realizabilities (cf. Section 4) during the duration of the step can be computed by classical models, e.g. using differential equations (in terms of the activities of the cat-neurons and the strengths of the links, cf. (Ehresmann and Vanbremeersch, 2009)). The result is evaluated and memorized at the beginning of the next step; there is a *fracture* for CR if the outcome is not the expected one, modeled by the fact that the next landscape (beginning at *t*′) is not the expected one and/or there is no admissible procept in it.

(ii) The global dynamic

The various co-regulators have different rhythms. At a given time the commands they sent to activate effectors (via their neuronal bases) can be conflicting; for instance to seize an object, the visual and motor commands should fit together. So there is need for an equilibration process, the *interplay among co-regulators*, to search for a best compromise between them, possibly causing fractures to them (e.g. by inhibition of their commands). An important role is played by *evaluating co-regulators*, based on parts of the emotive brain, which model the consequences on the well-being. Though the one-step local dynamics of the co-regulators could be computed by usual mathematical models, the interplay does not seem open to 'classical' computations. Indeed, it has a large number of freedom degrees since each effector can be activated through anyone of its neuronal realizabilities; moreover it must respect the structural temporal constraints of each co-regulator CR_i expressed by inequalities of the form $p_i \leq d_i \leq s_i$ which relate the length d_i of its step at t to the largest propagation delay p_i of the components of its landscape

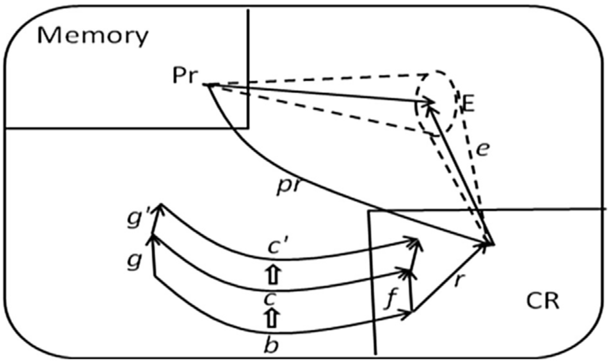

Fig. 6. The landscape of CR has for components the active (curved) links *b, c, c′, pr*. The links between them are commutative squares (with vertical composition).

and to their minimal stability-span s_i (Ehresmann and Vanbremeersch, 2007).

6. The Archetypal Core at the basis of phenomenological processes

The Structural Core of the human cortex (Section 5.1) plays a central integrative role in the dynamic of NEUR. The *Archetypal Core* (Ehresmann and Vanbremeersch, 2007) plays a similar role in MENS.

6.1. The Archetypal Core

Higher mental processes heavily depend on the formation of a repertory of memories of notable events and major recurrent experiences which integrate multiple modalities (sensorial, proprioceptive, motor,…) with their emotional overtones, their welfare values and the main procedures associated to them. Thanks to their frequent recall, they become connected by fast and strong links maintaining their activation for some time… With these links they form a particular sub-system of the memory which has been called the *Archetypal Core* (Ehresmann and Vanbremeersch, 2001, 2007); by associating sensory-motor and perceptual categories with value states, it represents a personal internal model of the body, its experiences and its acquired knowledge. The archetypal core is related to what Edelman calls the *value-dominated memory*: "a conceptually based special memory system for value matched to past categories" (Edelman, 1989, pp. 99–100).

In MENS, the Archetypal Core is modeled by a higher order integrative evolutive sub-system AC of the memory MEM with its neuronal base in the Structural Core (cf. 5.1). The records in AC bind patterns of records of different modalities, so that they are records of higher complexity order, with a number of ramifications with their base in SC, that is to say, a model in NEUR of the Structural Core. Its development over time depends on the emergence of records of increasing complexity order through iterated complexifications (made possible by the Emergence Theorem).

Due to the strongly connected structure of the SC, these 'archetypal records' are connected by multiple complex links. As they are often activated, with time these links become stronger and faster (generalized Hebb rule), and they form *archetypal loops* which propagate very quickly the activation of an archetypal record back to itself, thus self-maintaining it. Thanks to these loops, AC self-maintains its activity for some time.

An archetypal record A displays two interacting modes of temporal presence, conferring it a "duration" ("durée vraie" in the sense of Bergson, 1888). On the one hand, due to the long activation delay of a cat-neuron of higher order (cf. Section 4), its activation at the instant t implies that A has a ramification R whose neuronal base has been activated before t, thus A collects information coming from the just-passed through the unfolding of R. On the other hand, the activation of A can be temporally self-maintained through archetypal loops, so that A will remain activated later (because of the propagation delays of archetypal links), thus it also offers an opening to the next future. This situation compares to (and allows for) the *retention and protention* processes described by Husserl (1904): "At each moment, there is simultaneously in consciousness, the actual presence of phenomena already passed together with the anticipation or projection of the future".

6.2. The intentional network and its macro-landscape *ML*

AC is directly linked to some co-regulators CR_i of higher orders, based on what Crick (1995) calls *conscious units* (in associative brain areas). These co-regulators with the links connecting their components in MENS form an evolutive sub-system of MENS which we call the *intentional network* IN (to suggest its phenomenological analog). By acting as a macro-co-regulator, IN plays a main role for developing higher order processes in MENS, with AC acting as a conductor and a driving force.

A non-expected or arousing situation leads to a surge of attention correlated with the activation of part of the structural core. In MENS, archetypal records A,A′ which have a ramification R with its neuronal base in this part, are activated; their activation diffuses into AC via archetypal loops, then resonates to lower level cat-neurons via the unfolding of other ramifications of A and switches between them. It follows that a large domain of MENS is activated, and this activation at various levels persists for some time, thanks to the activation-maintaining role of AC (due to the archetypal loops). In particular this surging activity may cause a fracture to IN at t.

To counteract the fracture, IN acts as a macro-co-regulator to form a *macro-landscape* ML. It is an evolutive system on a longer timeline J than the timelines J_i of the landscapes L_i of the co-regulators CR_i; it has for components all the links activating IN during J, so that it unites and extends the L_i's both in 'depth' and in duration. Its development over time relies on the following processes:

(i) Sharing of information by the CR_i's: If b activates the component O of CR_i and f is an active link from O to an O′ in another CR_j then the composite fb is also a component of ML.
(ii) Access to past lower levels: If O is activated at t via a link a from an archetypal record A to O, it means A was activated at t-d_a (where d_a is the propagation delay of a). Due to the activation delay of a higher order cat-neuron, it means that the neuronal base of a ramification R of A has been activated still earlier, say at t-d_a-d'_a. Moreover we have components of ML (such as $c = ua$ and $b = vua$ on Fig. 7) coming from different levels of R, which have been activated at past times; these components (which did not figure in earlier landscapes of the CR_i) may model 'non-conscious' information. Thus ML conjugates a descent in the "depth of time" (in the sense of Vrobel, 2015) with a descent into lower levels (Fig. 8).
(iii) Propagation of the activation: The activation spreads via loops in AC to other archetypal records such as A′ and to their decompositions, leading to new components of ML such as a'' (cf. Fig. 7). Switches between decompositions of activated archetypal records also lead to new components (such as c').

ML has a longer timeline than the L_i's because its activation is

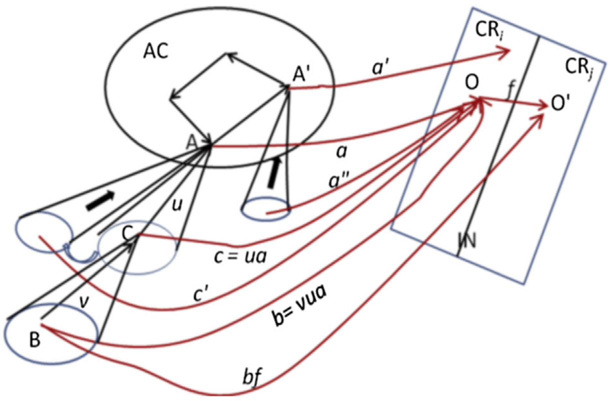

Fig. 7. AC, IN and formation of the macro-landscape ML. Its components are red.

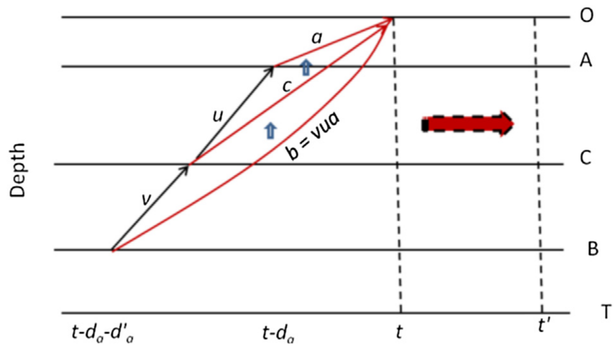

Fig. 8. Temporal origin of the components (in red) of ML in relation with their 'depth'. The large arrow indicates that the timeline of ML extends up to t'.

mediated by the activation of archetypal records and we know that AC is able to self-maintain its activity for some time. Thus, ML develops with time by iteration of the 3 preceding processes, its components coming from a range of levels and being activated at different times. (ML could be compared to the "theatre" of Baars, 1997.)

6.3. Higher cognitive processes

Successive temporally overlapping macro-landscapes give a frame for the development of higher cognitive and mental processes, such as conscious processes, embodied intelligence, creativity or anticipation (Ehresmann, 2012). It is done through iteration of the following intertwining processes.

(i) *Retrospection* (toward the past): Such a process starts by a non-expected or arousing situation ('phenomenon') which causes a fracture to the intentional network IN. As explained above, this leads to the formation of a macro-landscape ML which represents a dynamic model, for IN, of the system, its near past and its present becoming ('1st person' perspective of IN). Indeed, as we have seen, a 'temporal and depth descent', through unfolding ramifications, allows ML to retrieve past lower level events (for instance modeling instinctive motor or perceptive behaviors, emotions and affects, reflexes, …) thus accounting for embodiment and emotions. ML also contains higher levels information, and its various components can be shared between the 'conscious' CR$_i$, giving a comprehensive analysis of the situation.

(ii) *Prospection* (toward the future). It consists in the search for procedures on ML and their evaluation by construction of the corresponding complexification of ML; the long persistence of ML allows for realizing these operations; finally one is selected. Different methods are possible to construct such procedures. In particular, ML, among its components, contains components issued from activated archetypal procepts (say, a' in Fig. 7), for instance from admissible procepts of the CR$_i$'s. A procedure on ML can consist in binding patterns of such ML-components. The activation of the procepts is transmitted to their effectors, which may or not be seen in ML. In the first case, the corresponding complexification can be 'virtually' constructed in ML. Otherwise the complexification is constructed in MENS, and evaluated only through the changes caused to ML. These changes (e.g. the emergence of a complex link) may cause a fracture to ML, leading to the formation of a new macro-landscape and iteration of the retrospection and prospection processes in it.

The development of a higher cognitive process, say a creative or anticipative process, may require such a sequence of retrospection-prospection-complexification, with, at each step, selection of a procedure Pr_k on a macro-landscape ML_k and realization of the corresponding complexification; the sequence (Pr_k) of procedures is called a "scenario". The Iterated Complexification Theorem (Section 3) implies that, if complex links emerge in the intermediate complexifications, then the final result was not initially predictable (i.e. it cannot be obtained by a unique complexification of the first ML_0). It is the case with really creative scenarios, transcending the current situation with unpredictable results; examples are given by the "transformational creativity" of Boden (2004) or the "really new futures" of R. Miller (2007).

MENS has been constructed as a dynamic model of the neuro-cognitive mental system. Now we are going back to this system by trying to 'metaphorically' translate the model in terms of mental objects/processes/states, and their phenomenological meaning. IN becomes the conscious Self, ML its 1st person perspective, partially accessible to language.

The timeline of ML reflects the time of the "conscious acts themselves" (second level of temporality for Varela, 1999); it extends the physical time imposed by propagation and activation delays (his first level); while the overlapping of the timelines of successive macro-landscapes gives rise to the "flow of time related to personal identity" (his third level).

Simeonov (2015) writes: "… first is the question of how experiences at large could already be latent in time experiences of the individual. Once this has been clarified, the next step should be to identify the nature of an agency for the succeeding abstraction along Mach's line. Maybe the notion of delays could bridge these two projections." His suggestion is somewhat justified by the specific role played by the propagation delays and activation delays in the AC and ML operations (as explained in 6.2 and 6.3).

The retrospection process consists in the formation and analysis of ML for 'making sense' of the present situation. Prospection is an intentional process to select and evaluate long-term scenarios for the future (while protention just indicates an opening toward the future, a "disposition for action" (Varela, 1999)). The construction of scenarios can be compared to the construction of *narratives* in (Goranson and Devlin, 2015).

7. Discussion and conclusion: MENS vs. classicism, connexionism and neuro-phenomenology

MENS gives a dynamic mathematical methodology for studying the neuro-cognitive-mental system as a whole. Using categorical constructions (colimits, HES, complexification process), it explains how a hierarchy of mental objects emerges from brain activity, a mental object having a dynamic multi-faceted representation, under the form of a category-neuron which takes its own individuation over time. A key factor for their emergence and multiplicity is the *degeneracy* property of the neural code (formalized by the *Multiplicity Principle*). By combination of such cat-neurons depending on certain rules (represented by the links of the patterns that they bind), mental representations (objects and processes) and mental states of increasing complexity can emerge and allow developing a flexible memory, with a central part, the *Archetypal Core* AC which acts as a conductor for the formation of macro-landscapes in which higher order cognitive and mental processes can be developed.

Thus, MENS seems to set the basis for a truly multidisciplinary approach to higher order brain function and mental processes, and to the mind—brain interaction problem.

To conclude we address the 2 following questions:

(i) To what kind of 'Theory of Mind' does MENS lead? in particular, how is it situated in relation to classicism and connectionism?
(ii) Is MENS entitled to be called a 'Naturalized Phenomenology'?

7.1. MENS as a theory of mind between classicism and connectionism

The Stanford Encyclopedia of philosophy defines a "*Representational Theory of Mind*" as "any theory that postulates the existence of semantically evaluable mental objects, <…> as well as the various sorts of "subpersonal" representations postulated by cognitive science ". In particular a Theory of Mind should satisfy two main properties, namely *compositionality* and *productivity*.

These properties are satisfied in MENS: cat-neurons are obtained by 'composition' (binding) of (hyper-)assemblies of neurons; but they themselves can be combined to 'produce' new cat-neurons which, thanks to MP, can represent more and more complex mental states, whence the compositionality and productivity of mental processes represented by cat-neurons.

These two properties are also satisfied in classicism, but in it, at the difference of MENS, the representations are symbolic and based on 'atoms', while cat-neurons are multi-faceted dynamic representations, taking their own individuation over time, and with an internal structure with multiple neuronal realizabilities. In connectionism, representations may have a neural basis as for us, but Fodor and Pylyshyn (Fodor and Pylyshyn, 1988) have shown that connectionist models cannot handle compositional semantics. In MENS it becomes possible because we introduce the 'binding' operation, leading to a kind of hierarchy of connectionist systems, each one having for units at its base the cat-neurons of strictly lower levels.

Thus, MENS seems intermediate between classicism and connectionism, and somewhat closer to some neo-connectionism models such as morphodynamism in which

> " the conceptual contents of mental states and their semantic correlates are no longer identified with labels for symbolic ordering. Their meaning is embodied in the cognitive processing itself, identified with the topology of complex attractors of the underlying neural dynamics, and the mental events are identified with sequences of bifurcations of such attractors" (Barandiaran, 2011).

Indeed, it can be proved (Ehresmann and Vanbremeersch, 2009) that locally a cat-neuron plays the role of an attractor of the neural dynamics; and a cat-neuron, being multi-faceted, keeps its individuation over time by addition or suppression of some ramifications (acting as bifurcations of its trajectory).

7.2. Systematicity in MENS

Another property generally required for a theory of mind is *systematicity*. To analyze if it is satisfied in MENS, we first need to say what it means. The problem is that there is currently no encompassing definition which everyone agrees upon.

In the Dictionary of Philosophy of Mind, systematicity is defined as a number of putative psychological properties or regularities. Fodor and Pylyshyn define systematicity of mental representation as "the fact that cognitive capacities always exhibit certain symmetries, so that the ability to entertain a given thought implies the ability to entertain thoughts with semantically related contents. " More explicitly, for (Hadley, 1997):

> " A cognitive agent, *C*, exhibits systematicity just in case its cognitive architecture causally ensures that *C* has the capacity to be in a propositional attitude (*A*) towards proposition *aRb* if and only if *C* has the capacity to be in attitude (*A*) towards proposition *bRa*.

Similar ideas have been elaborated upon at length elsewhere, in particular in Kenneth Aizawa's book (Aizawa, 2003).

These definitions essentially concern context-independent capabilities, but even in the case of language there are exceptions, for instance $x \in \{x\}$ does not allow $\{x\} \in x$ in the usual language of set theory including the foundation axiom. This has led some authors, e.g. Johnson (Johnson, 2004), to refute any kind of systematicity. A more formal approach using category theory has been developed by Phillips and Wilson (Phillips and Wilson, 2010), (Phillips and Wilson, 2011); for them systematic (and quasi-systematic) properties are instances of universal constructions, and may, in particular, involve the capacity to compute coproducts.

In MENS, we propose 2 kinds of systematicity:

(i) *Bottom-up systematicity*: As indicated in Sections 3 and 4, cat-neurons (constructed as colimits) and simple links depend on *systematic constructions* based on universal properties; and complex links can be interpreted as kinds of new systematic rules imposed by the categorical composition.
(ii) *Top-down systematicity*: We have seen (Section 4.4) that a cat-neuron has multiple neuronal realizabilities through the activation of its different ramifications. It follows that the capacity of activating one of these ramifications systematically implies the capacity of activating the others.

7.3. MENS provides a naturalized phenomenology

As a consequence of the MP, a mental object, represented by a multi-faceted cat-neuron, has multiple physical realizations into hyper-assemblies of neurons (through its different ramifications), thus allowing for flexibility and 'multiple' mental causation. This multiplicity shows that Varela's Hypothesis 1:

> "For every cognitive act, there is a singular, specific cell assembly that underlies its emergence and operation" (Varela, 1999)

is only valid for tokens, not for types, and it explains how MENS avoids the "isomorphism between neural and mental" at the basis of many criticisms of neuro-phenomenology.

Please cite this article in press as: Ehresmann, A.C., Gomez-Ramirez, J., Conciliating neuroscience and phenomenology via category theory, Progress in Biophysics and Molecular Biology (2015), http://dx.doi.org/10.1016/j.pbiomolbio.2015.07.004

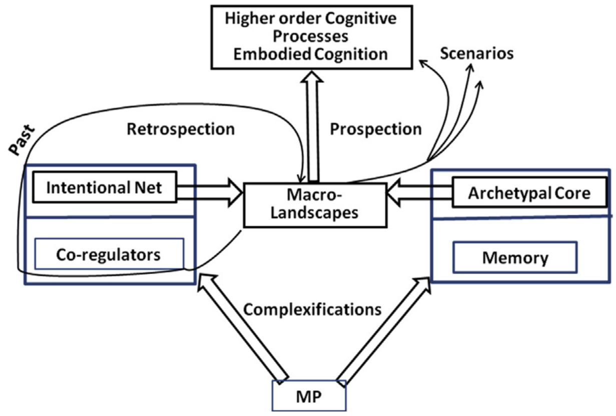

Fig. 9. The figure outlines the main notions introduced in MENS which allow for the development of higher order cognitive processes.

To summarize, MENS proposes a categorical "bridging strategy" (Bayne, 2004) for closing the gap between neural and phenomenal data by also accounting for intentionality and consciousness, Bergsonian duration, Husserl's retention and protention, embodied cognition, higher order cognitive and mental processes up to creativity and anticipation. It is done via the long term macro-landscapes ML, which model the "1st person" perspective; their construction and duration depend on the Archetypal Core (the mental analog of the Structural Core of the brain) which acts as a driving motor by self-maintaining its activation for some time through its archetypal loops and transmitting it to ML (Fig. 9).

Appendix. Mathematical definitions

1 An *Evolutive System* K consists of:
  (i) The timeline T of the system and, for each t of T, a category K_t called *configuration* of the system at t.
  (ii) For each time $t' > t$, a functor $k_{t,t'}$ *transition* from t to t' from a subcategory of K_t to $K_{t'}$. These functors satisfy the transitivity condition:

(TC) If an object A_t has new configuration $A_{t'} = k_{tt'}(A_t)$ at t', then $A_{t'}$ has a new configuration $A_{t''} = k_{tt''}(A_t)$ at t'' if, and only if, $A_{t'}$ has a configuration at t'', and then $A_{t''} = k_{t't''}(A_{t'})$. Similarly for the links.

2. If P and P′ are two patterns in a category C, a *cluster* from P to P′ is a maximal set G of links between their components satisfying the following conditions:
  (i) Each P_i has at least one link to a component of P′; and if there are several such links, they are correlated by a zigzag of distinguished links of P′.
  (ii) The composite of a link in G with a distinguished link of P′, or of a distinguished link of P with a link in G also belongs to G.

If P and P′ have colimits M and M′ respectively, it follows from the universal property of a colimit that the cluster G 'binds' into a unique link cG from M to M′, called a (P, P′)-*simple link.*

Two patterns Q and P are *non-connected* if they have the same colimit M though there is no cluster between them binding into the identity of M. Then M is said to be *multi-faceted*, and the passage from Q to P a *switch* between decompositions of M.

3. *Complexification*: A *pro-sketch* Pr on a category K consists (at least) of the following data: a set S of objects of K and a set Π of patterns in K. The complexification K′ or K with respect to Pr is the 'universal solution' of the problem of constructing a category K′, called the *complexification of* K for Pr, and a functor F from a sub-category K° of K to K′ verifying the conditions:
  (i) K° does not contain the elements of S;
  (ii) for each P in Π the pattern FP image of P by F admits a colimit cP in K′; if P has a colimit in K, cP is the image of this colimit by F; otherwise cP 'emerges' in K′ to become the colimit of P. The morphisms of K′ are both simple links and complex links.

For the explicit construction of the complexification, cf. EV (1987, 2007).

References

Aizawa, Kenneth, 2003. The Systematicity Arguments. Kluwer Academic, Boston.
Awodey, S., Gambino, N., Lumsdaine, P.L., Warren, M.A., 2009. Lawvere—Tierney sheaves in algebraic set theory. J. Symbolic Log. 74 (3), 861–890. http://dx.doi.org/10.2178/jsl/1245158088.
Aydede, Murat, Fall 2010. The language of thought hypothesis. In: Zalta, Edward N. (Ed.), The Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/archives/fall2010/entries/language-thought/.
Baars, Bernard J., 1997. In the Theater of Consciousness: The Workspace of the Mind. Oxford University Press, New York.
Barandiaran, Xabier, 2011. From systematicity of thought to systematicity of habits. In: Systematicity and the Post-Connectionist Era Workshop, San José, Spain.
Barlow, H.B., 1972. Single units and sensation: a neuron doctrine for perceptual psychology? Perception 1 (4), 371–394. http://dx.doi.org/10.1068/p010371.
Bayne, Tim, 2004. Closing the gap? Some questions for neurophenomenology. In: Phenomenology and the Cognitive Science3/4, pp. 349–364.
Bergson, Henri, 1888. Essai sur les données immédiates de la conscience. In: Bergson, Henri (Ed.), Oeuvres P.U.F. 1959.
Boden, Margaret, 2004. The Creative Mind; Myths and Mechanisms, second ed. Routledge.
Borceux, Francis, 2009. Handbook of Categorical Algebra. In: Categories and Structures, vol. 2. Cambridge University Press, Cambridge.
Bullmore, Ed, Sporns, Olaf, 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10 (3), 186–198. http://dx.doi.org/10.1038/nrn2575.
Buzsaki, György, 2010. Neuronal syntax: cell assemblies, synapsembles, and readers. Neuron 68 (3), 362–385. http://dx.doi.org/10.1016/j.neuron.2010.09.023.
Crick, Francis, 1995. The Astonishing Hypothesis: the Scientific Search for the Soul. Simon & Schuster, New York.
Destexhe, Alain, 2011. Intracellular and computational evidence for a dominant role of internal network activity in cortical computations. Curr. Opin. Neurobiol. 21 (5), 717–725. http://dx.doi.org/10.1016/j.conb.2011.06.002.
Edelman, Gerald M., 1989. The Remembered Present: a Biological Theory of Consciousness. Basic Books, New York.
Ehresmann, Charles, 1957. Gattungen von lokalen Strukturen. Jahresber. Dtsch. Mathematiker-Vereinigung 60, 49–77.
Ehresmann, Andrée C., 2012. MENS, an info-computational model for (neuro-) cognitive systems capable of creativity. Entropy 14 (12), 1703–1716. http://dx.doi.org/10.3390/e14091703.
Ehresmann, A.C., Vanbremeersch, J.-P., 1987. Hierarchical evolutive systems: a mathematical model for complex systems. Bull. Math. Biol. 49 (1), 13–50. http://dx.doi.org/10.1007/BF02459958.
Ehresmann, A.C., Vanbremeersch, Jen-Paul, 1996. Multiplicity principle and emergence in MES. J. Syst. Anal. Model. Simul. 26, 81–117.
Ehresmann, A.C., Vanbremeersch, Jean-Paul, 2001. Emergence processes up to consciousness using the multiplicity principle and quantum physics. In: Dubois, D. (Ed.), ALP. Conference Proceedings, 627, pp. 221–233. CASYS.
Ehresmann, A.C., Vanbremeersch, J.P., 2007. Memory Evolutive Systems; Hierarchy, Emergence, Cognition, first ed., vol. 4. Elsevier Science.
Ehresmann, A.C., Vanbremeersch, J.-P., 2009. A propos des Systèmes Evolutifs à Mémoire et du modèle MENS. In: Compte-rendu du SIC (Séminaire Intinérant des Catégories), Paris.
Eilenberg, Samuel, MacLane, Saunders, 1945. General theory of natural equivalences. Trans. Am. Math. Soc. 58 (2), 231. http://dx.doi.org/10.2307/1990284.
Fodor, J.A., Pylyshyn, Z.W., 1988. Connectionism and cognitive architecture: a critical analysis. Cognition 28 (1–2), 3–71.
Freeman, Walter J., Vitiello, Giuseppe, 2006. Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. Phys. Life Rev. 3 (2), 93–118. http://dx.doi.org/10.1016/j.plrev.2006.02.001.
Friston, Karl, 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11 (2), 127–138. http://dx.doi.org/10.1038/nrn2787.
Gatherer, Derek, Galpin, Vashti, 2013. Rosen's (M,R) system in process algebra. BMC Syst. Biol. 7 (1), 128. http://dx.doi.org/10.1186/1752-0509-7-128.
Gomez, Jaime, Sanz, Ricardo, 2009. To cognize is to categorize revisited: category theory is where mathematics meets biology. In: 2009 AAAI Fall Symposium

Please cite this article in press as: Ehresmann, A.C., Gomez-Ramirez, J., Conciliating neuroscience and phenomenology via category theory, Progress in Biophysics and Molecular Biology (2015), http://dx.doi.org/10.1016/j.pbiomolbio.2015.07.004

Series. http://aaai.org/ocs/index.php/FSS/FSS09/paper/view/964.
Gomez-Ramirez, Jaime, 2014. A New Foundation for Representation in Cognitive and Brain Science – Category Theory and the Hippocampus. In: Springer Series in Cognitive and Neuronal Systems, 7. Springer, Netherlands. http://www.springer.com/biomed/neuroscience/book/978-94-007-7737-8.
Gómez-Ramirez, Jaime, Sanz, Ricardo, 2011. Hippocampal categories: a mathematical foundation for navigation and memory. In: Hernández, Carlos, Sanz, Ricardo, Gómez-Ramirez, Jaime, Smith, Leslie S., Hussain, Amir, Chella, Antonio, Aleksander, Igor (Eds.), From Brains to Systems, 718. Springer New York, New York, NY, pp. 149–164. http://www.springerlink.com/content/k3m12613r71158jl/.
Gómez-Ramirez, Jaime, Sanz, Ricardo, 2013. On the limitations of standard statistical modeling in biological systems: a full Bayesian approach for biology. Prog. Biophys. Mol. Biol. 113 (1), 80–91. http://dx.doi.org/10.1016/j.pbiomolbio.2013.03.008.
Goranson, Ted, Devlin, Keith, 2015. Pragmatic phenomenological types. J. Prog. Biophys. Mol. Biol. Vol. XXX, Issue Y. Special Theme Issue on Integral Biomathics: Life Sciences, Mathematics, and Phenomenological Philosophy. Elsevier. ISSN: 0079–6107. (in this issue).
Grothendieck, Alexander, 1957. Sur Quelques points D'algèbre Homologique, I. Tohoku Math. J. 9 (2), 119–221. http://dx.doi.org/10.2748/tmj/1178244839.
Hadley, Robert F., 1997. Cognition, systematicity and nomic necessity. Mind Lang. 12 (2), 137–153. http://dx.doi.org/10.1111/j.1468-0017.1997.tb00066.x.
Hagmann, Patric, Cammoun, Leila, Gigandet, Xavier, Meuli, Reto, Honey, Christopher J., Wedeen, Van J., Sporns, Olaf, 2008. Mapping the structural core of human cerebral cortex. PLoS Biol. 6 (7), e159. http://dx.doi.org/10.1371/journal.pbio.0060159.
Hebb, David, 1949. The Organization of Behaviour. Wiley, New York.
Hubel, D.H., Wiesel, T.N., 1977. Ferrier lecture. Functional architecture of Macaque Monkey visual cortex. Proc. R. Soc. Lond. Ser. B, Contain. Pap. a Biol. Charact. R. Soc. (Great Britain) 198 (1130), 1–59.
Izhikevich, Eugene M., 2006. Dynamical Systems in Neuroscience: the Geometry of Excitability and Bursting, first ed. The MIT Press.
Izhikevich, Eugene M., Gally, Joseph A., Edelman, Gerald M., 2004. Spike-timing dynamics of neuronal groups. Cereb. Cortex (New York, N.Y. 1991) 14 (8), 933–944. http://dx.doi.org/10.1093/cercor/bhh053.
Johnson, Kent, 2004. On the systematicity of language and thought. J. Philos. 101 (3), 111–139.
Kan, Daniel M., 1958. Adjoint functors. Trans. Am. Math. Soc. 87 (2), 294. http://dx.doi.org/10.2307/1993102.
Kauffman, S., Gare, A., 2015. Beyond descartes and Newton: recovering life and humanity. J. Prog. Biophys. Mol. Biol. Vol. XXX, Issue Y. Special Theme Issue on Integral Biomathics: Life Sciences, Mathematics, and Phenomenological Philosophy. Elsevier. ISSN: 0079–6107 (in this issue).
Kim, Jaegwon, 1998. Mind in a Physical World: an Essay on the Mind-body Problem and Mental Causation. MIT Press, Cambridge, Mass.
Landry, Elaine, 1999. Category theory: the language of mathematics. Philos. Sci. 66 (3), 27.
MacLane, Saunders, 2010. Categories for the Working Mathematician. Softcover reprint of hardcover, second ed. Springer. 1998.
Letelier, Juan-Carlos, Soto-Andrade, Jorge, Guíñez Abarzúa, Flavio, Cornish-Bowden, Athel, Luz Cárdenas, María, 2006. Organizational invariance and metabolic closure: analysis in terms of (M,R) systems. J. Theor. Biol. 238 (4), 949–961. http://dx.doi.org/10.1016/j.jtbi.2005.07.007.
Malsburg, Christoph von der, Bienenstock, Elie, 1986. Statistical coding and short-term synaptic plasticity: a scheme for knowledge representation in the brain. In: Bienenstock, E., Fogelman Soulié, F., Weisbuch, G. (Eds.), Disordered Systems and Biological Organization, pp. 247–272. NATO ASI Series 20. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-82657-3_26.
Miller, Riel, 2007. Futures literacy: a hybrid strategic scenario method. Futures 39.
Phillips, Steven, Wilson, William H., 2010. Categorial compositionality: a category theory explanation for the systematicity of human cognition. PLoS Comput. Biol. 6 (7), e1000858. http://dx.doi.org/10.1371/journal.pcbi.1000858.
Phillips, Steven, Wilson, William H., 2011. Categorial compositionality II: universal constructions and a general theory of (Quasi-)Systematicity in human cognition. PLoS Comput. Biol. 7 (8), e1002102. http://dx.doi.org/10.1371/journal.pcbi.1002102#references.
Popper, K., 1972. Objective Knowledge, an Evolutionary Approach. Clarendon Press, Oxford.
Rosen, Robert, 1958. The representation of biological systems from the standpoint of the theory of categories,. Bull. Math. Biophys. 20, 317–341.
Rosen, Robert, 1985b. Organisms as causal systems which are not mechanisms. In: Theoretical Biology and Complexity. Acad. Press, New York, pp. 165–203.
Rosen, Robert, 1985a. Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations. Pergamon Press.
Simeonov, Plamen, 2015. Yet another time about time…part I: an essay on the phenomenology of physical time. J. Prog. Biophys. Mol. Biol. Vol. XXX, Issue Y. Special Theme Issue on Integral Biomathics: Life Sciences, Mathematics, and Phenomenological Philosophy. Elsevier. ISSN: 0079–6107 (in this issue).
Spivak, David, 2014. Category Theory for the Sciences. MIT Press.
Thom, René, 1988. Esquisse d'une sémiophysique. InterÉditions, Paris.
Varela, Francisco, 1996. Neurophenomenology. J. Consciousness Stud. 3, 330–350.
Varela, Francisco, 1999. The specious present: a neurophenomenology of time consciousness. In: Naturalizing Phenomenology. Stanford Univ. Press, Stanford, pp. 266–329. Chapter 9.
Vrobel, Susie, 2015. A new kind of relativity: compensated delays as phenomenal blind spots. J. Prog. Biophys. Mol. Biol. Vol. XXX, Issue Y. Special Theme Issue on Integral Biomathics: Life Sciences, Mathematics, and Phenomenological Philosophy. Elsevier. ISSN: 0079–6107 (in this issue).

Please cite this article in press as: Ehresmann, A.C., Gomez-Ramirez, J., Conciliating neuroscience and phenomenology via category theory, Progress in Biophysics and Molecular Biology (2015), http://dx.doi.org/10.1016/j.pbiomolbio.2015.07.004