Progress in Biophysics and Molecular Biology 119 (2015) 347–359

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

ARTICLE INFO

Article history:
Available online 17 July 2015

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

ABSTRACT

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/

(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 neuro-phenomenology 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 neuro-phnomenology.

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

$$(f_1, f_2, \ldots, f_n) \text{ with } f_1 : \mathrm{A} \to \mathrm{A}_1 \cdot f_2 : \mathrm{A}_1 \to \mathrm{A}_2, \ldots, f_n : \mathrm{A}_{n-1} \to \mathrm{B}.$$

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.