

A formal approach of developmental robotics and psychology

Ken Prepin* **
 prepin@ensea.fr

Philippe Gaussier*
 gaussier@ensea.fr

Arnaud Revel*
 revel@ensea.fr

Jacqueline Nadel**
 jnadel@ext.jussieu.fr

* Neuro-cybernetic team, ETIS
 UCP / ENSEA
 95014 Cergy Pontoise, France

**Vulnérability, Adaptation, Psychopathology lab,
 La Salpêtrière Hospital, Paris
 750013 Paris, France

Four years ago, we proposed a Cognitive System Formalism (CSF) to represent, analyze and compare cognitive systems (2). The building of this formalism is an attempt to bridge the gap between theories on dynamical systems, long life learning and experiments in robotics and in developmental psychology. The formalism is based on the deep structural link that exists between Perception and Action in any autonomous and embodied “intelligent” system. Each Perception is defined as a potential function corresponding to a behavioral attraction basin (1, 7). Consequently we define the Action as the gradient of this Perception $\vec{Ac} = -m.grad(Per)$ (3, 5).

In the formalism, two types of simplification rules were introduced. Simplifications of the first type can be performed at any time and leave the fundamental properties of the system completely unchanged. These are very limited simplification rules which only allow to reduce the number of functional boxes in the architecture and to suppress trivial redundancies (simplification of a chain of reflex links). Simplifications of the second type are more interesting but can only be applied after learning is stabilized (i.e. the system is in a given state of Perception). They allow to formalize easily the interaction with the environment or another agent and to characterize the importance of the embodiment. After simplification, the resulting system should be equivalent to the original even if it is a different system (possibly less adaptive and robust to environmental changes). Hence, the choice of formal simplification rules implies to explicit what is fundamental in cognitive processes. These choices should be criticized by the community in order to define a coherent set of fundamental equations taking into account a larger and more precise range of situations (i.e. defining new invariants linked to principles not taken into account in the current version of our formalism).

From an applicative point of view the formalism

must allow to analyze “intelligent” systems considering the nature of the data flows between their different decision and/or memory elements in order to predict their stable states and their dynamical properties.

In our formalism, a CS is supposed to be made of several elements or nodes or boxes associated with input information, intermediate processes and output (command of actions). Functional links between boxes are represented by matrices. Two main types of W matrix are distinguished according to their learning capabilities (learning possible or not): *Unconditional* matrix U for reflex mechanisms and *Adaptive* matrix A which are used for pattern matching processes, categorization... or all the other ways of filtering performed by learning. Each box is characterized by an operator k . It defines both the way to use the weight matrices in order to compute the output and the way to modify the matrices according to a given learning criteria. The operator can include a non-linear function and a pattern of interaction between the elements of the same block: lateral interactions in the case of a competitive structure, recurrent feedback in the case of a dynamical system... The operator can be a more complex algorithmic element (any “if...then...else...” treatment can be expressed in terms of dynamical lateral interactions but the reciprocal is false). Our mathematical formalism is inspired from quantum mechanics. Indeed, both actions and sensory information can be seen either as atomic elements or oscillatory signals according to the tasks (for instance, actions can be modeled by a series of via points or at the opposite by a composition of oscillatory signals). Yet, in our case the space is not Hilbertian since the operators are not linear.

In previous works, our formalism has been used to explain how babies learn to associate seen but unfelt facial expressions of adults with their own felt but unseen facial expressions (3). We mathemati-

cally proved that a simple visuo-motor architecture coupled with a low level emotional system was sufficient to learn the association if the parents act as a mirror of their baby facial expressions. Predictions of the mathematical model are now tested with both psychological experiment and computational models. At the opposite, we have recently studied the robotic joint-attention architecture developed and tested by Hosoda (4). Our goal was to show that our formalism can be applied to analyze an architecture proposed by another team even if the computational tools are different from ours and not completely detailed in the papers. We summarize here the necessary steps of such study (detailed in (6)). First, the schematic architecture and the associated algorithms have been translated in our mathematical formalism. The equations represent both the flow of informations, the decisions making and the learning mechanisms used in the architecture fig 1. Next, to study

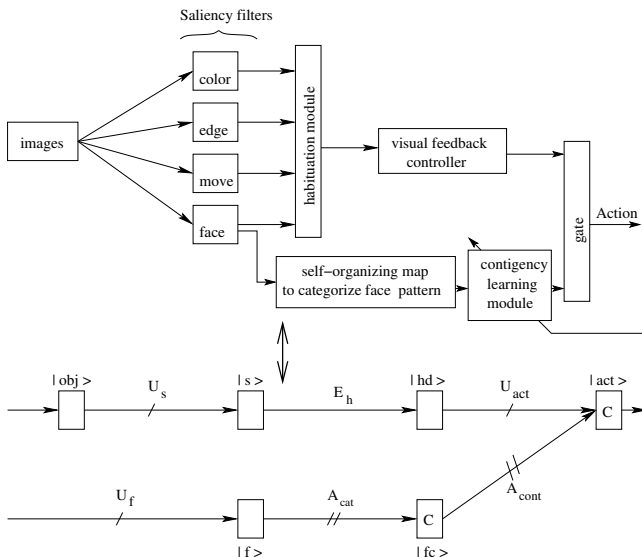


Figure 1: Reproduction of the joint attention architecture of Hosoda et al.: $|obj\rangle$ is the vector describing objects, U_s is the unconditional saliency matrix, in the same way h is for habituation, hd is for habituation decision, f is for face, A_{cat} is for the adaptive categorization matrix, fc is for face category, $cont$ is for contingency, act is for action, C is for competition mechanism. We can clearly see the pathway of the signal through the different modules.

the attractors of this dynamical system, we immerse it in its environment (figure 2). The formalism allows us to see clearly the possible recurrent aspect of the dynamics of the architecture. Here, the action of the studied agent 1: $|act_1\rangle$ depends on its own previous action through the operator k . This operator represents the way the second agent links the objects present in the environment to agent 1's actions. Our theoretical analysis allows us to predict and extend the Hosoda et al. results (showing what are the nec-

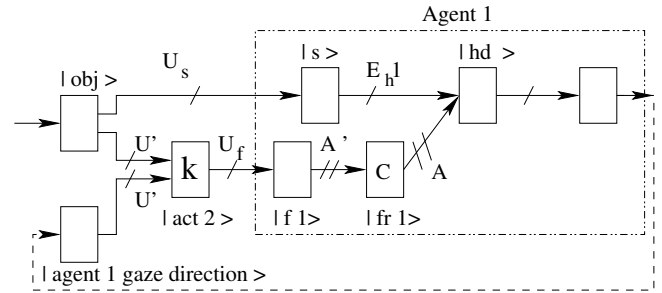


Figure 2: *Agent 2's* behavior depends on both objects and on *Agent 1* (the robot) behavior, through the operator k .

essary and the best conditions for the stabilization of learning). Finally, we have shown that a small change on the dynamics of A_{cont} (enlargement of the temporal window of contingency learning) modifies the recursive relation of the *sensation-action* loop and thus the development of the robot (6). In this case, the agent should be able to learn from being imitated: a powerful capability, leading to a faster convergence of learning (3). Future works and critics should allow to refine the formalism and to extend it to more complex problems such as long term development including a sequence of emergent stages and more complex dynamics of interaction (turn taking, role switching...).

References

- [1] R.D. Beer. A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, (72):173–215, 1995.
- [2] P. Gaussier. Toward a cognitive system algebra: A perception/action perspective. In *EWLR*, pages 88–100, http://www-etis.ensea.fr/~neurocyber/EWRL2001_gaussier.pdf, 2001.
- [3] P. Gaussier, J.C. Baccon, K. Prepin, J. Nadel, and L. Hafemeister. Formalization of recognition, affordances and learning in isolated or interacting animats. In *SAB'04*, MIT Press, Los Angeles, 2004.
- [4] K. Hosoda, H. Sumioka, A. Morita, and M. Asada. Acquisition of human-robot joint attention through real-time natural interaction. In *International Conference on Intelligent Robots and Systems*, volume 3, pages pp.2867–2872. IEEE/RSJ, 2004.
- [5] M. Maillard, O. Gapenne, L. Hafemeister, and P. Gaussier. Perception as a dynamical sensori-motor attraction basin. In *LNCS - ECAL05*, in press, Canterbury, 2005.
- [6] K. Prepin. *A formal approach of developmental robotics and psychology: application to the study of a joint attention architecture*, internal report. <http://www.etis.ensea.fr/~neurocyber/epirob2005>, 2005.
- [7] A. Revel and J. Nadel. *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions.*, chapter How to build an imitator? Cambridge University Press., to be published edition, 2004.