2012 IEEE International Conference on Systems, Man, and Cybernetics
October 14-17, 2012, COEX, Seoul, Korea

A model of a middle level of cognition based on the interaction among the thalamus, amygdala, and the cortex

Riccardo Manzotti and Flavio Mutti and Soo-Young Lee and Giuseppina Gini

Abstract— **How the cognitive capabilities of humans are mapped on the underlying computational framework is far to be completely understood. Despite it is an open problem, some evidences show that the interaction among the cortex, the amygdala and the thalamus is involved in the composition of elementary behaviours to solve complex tasks. The cortex is well-known to have a certain degree of plasticity to categorise the processed information whereas the other two brain areas are involved in the association of an emotional state to the incoming sensory information. In the specific, the amygdala is involved in the generation of innate affective states whereas the thalamus deals with the developing of new goals. The interaction among these three brain structures plays an important role in the emerging of a middle level of cognition. We claim that the cognitive model could fill the gap between the high-level reasoning and low level sensory-motor coordination. In this paper, firstly we model the above mentioned brain areas and second we investigate the interaction of these areas, producing a goal-generating architecture. Finally, the goal-generating system is tested on a NAO robot, in order to validate and extend some previously published results.**

I. Introduction

This paper aims to outline a cognitive architecture capable of addressing a level of cognition that has not received overwhelming attention so far. As it will get clearer in the following, we do not address fine sensorimotor skills nor high level logical reasoning. Rather we focus on that intermediate level of cognition that allows humans and mammals to interact with the environment for most of our daily routines. In 2006, this level has been the target of some attention by Ray Jackendoff who suggested that "The central hypothesis I arrive at is called the Intermediate-Level Theory of Consciousness: the form of awareness is derived neither from "sense-data" in the traditional sense nor from the form of thought, as many have guessed, but rather in each separate faculty of the mind-from a level of representation intermediate between the most peripheral (sensation-like) and the most central (thought-like)" [7]. Although we do not share Jackendoff's focus on linguistic and computational aspects, we share his interest to the existence of a cognitive level bridging the gap between high level conceptual reasoning and low level fine-grained sensory-motor control. An example may help to flesh out the matter. Consider an average human being: Homer. It is 6 a.m. Homer wakes up. He gets into the bathroom still half asleep. His actions are neither accurate nor the result of complex reasoning. He is simply following some vague goals and routine behaviour. Once he sees his razor, he remembers to shade. While he shades, he cuts himself. Then he pours cold water on his cut. In the shower, he showers muttering a funny tune he would not be able to name. All of a sudden, he sees his mobile phone abandoned on a shelf. He had no clear idea why it is there. Nevertheless, he grabs it for further use. And so forth. There is no explicit logical reasoning to drive his actions (although they show some purpose) and they are not the result of pure sensorimotor coordination. With this example we want to stress that most human actions are neither the result of sensorimotor contingencies nor the output of logical reasoning. Rather it is some intermediate quick-and-dirty cognitive level roughly corresponding with our conscious awareness. It allows being situated in the environment and shifting from one situation to the next while loosely pursuing one's goals. It is a level that it is fair to assume plays an important role in non-human mammals' cognitive life which cannot take advantage of symbolic logical reasoning. When an agent aims at something, the sensorimotor planning required is often outsourced to more specialized neural circuitry. This is the level we plan to address in this paper. Such an intermediate level is appealing since it might endorse a series of important features usually related with consciousness: semantics, purposefulness, choice. For instance, when we refer either to an object or a person, this happens in some intermediate level between pure visual processing and high level semantics. The so-called symbol grounding problem requires a world which is neither made of high-level concepts nor of fine-grained motor controls. Thus the intermediate level might be well suited for grounding concepts and symbols to world facts. Thus, we may temporarily consider cognition as made of three levels: a sensorimotor level that embeds the fine motor control and perceptual processing; an intermediate level roughly corresponding to daily awareness and a higher level where long-term planning and complex symbolic reasoning are carried on.

The main contributions of this paper are:

- definition of a new unsupervised learning method for developing new goals;
- introduction of the amygdala model in the overall architecture;
- validation of previous results with the new cognitive

This work has been possible thanks to the support provided by the 2010-2012 Italian-Korean bilateral project (ICT-CNR and KAIST).

F. Mutti and G. Gini are with Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy mutti@elet.polimi.it

R. Manzotti is with Institute of Communication and Behaviour, IULM University, Milano, Italy riccardo.manzotti@iulm.it

Soo-Young Lee is with Brain Science Research Center, KAIST, Daejeon, Republic of Korea

978-1-4673-1714-6/12/$31.00 ©2012 IEEE

features.

This paper is organized as follows. In Section II we present some related works, in Section III we design the goal-generating architecture, in Section IV we present the performed experiment and in Section V we draw our conclusions.

II. THALAMO-AMYGDALA-CORTICAL MODEL

The amygdala-thalamo-cortical model is focused on the interaction among these areas. In the following we describe those functionalities that are considered in our model and how the interaction is exploited.

A. Cortex

One of the main feature of the cortex is its universal capability to adapt to novel stimuli regardless their physical sources, e.g. visual or proprioceptive [2][4]. Thus, it is reasonable to investigate a general approach to stimuli representation that exploits its plasticity, albeit with some drawbacks (a trade-off between generality and efficiency). Experimental evidences show the cortex capability to exhibit an high degree of plasticity and cognitive autonomy, pointing out that each cortical area has the capability to implement any computational skill [13][14][15].

Some evidences clearly show how the statistics of the incoming stimuli plays an important role in emergent shape of the neurons receptive fields. In [12], Olshausen et al. point out how it is possible to infer the receptive fields of the primary visual cortex neurons by means of statistical tools. Moreover, Independent Component Analysis (ICA) is a candidate to be a general way of input representation regardless the kind of stimuli [5][6].

For the purposes of our work and without loss of generality, we will implement a slightly simpler algorithm of input representation, namely a classical clustering algorithm, because as first step the aim of the project is to produce a proof-of-concept of the overall architecture capabilities.

B. Thalamus

As to goal generation, it is well known that most mammals are capable not only to learn how to achieve goals but also to learn what goals have to be pursued [8]. Putting together generality and goal generation may lead to suspect that goal generation is indeed spread everywhere in the cortex and that it is achieved by means of the interplay between thalamus and cortex. As Ward outlined in [16], "The thalamus is a "miniature map" of the cortex [...] One of the most compelling reasons why the thalamus has figured prominently in theories of consciousness is that it represents a central, convergent, compact "miniature map" of the cortex". Thus the thalamus is closely coupled with the cortex and, since most of data processing is carried on in the cortex itself, such interplay must have a different explanation. We suspect that each partition in the thalamus provides, to the corresponding cortical area, control information regarding what has to be pursued while the cortex will keep processing data [10]. After all, it is well known that the more massive cortico-thalamic projections are those proceeding from the cortex to the thalamus [11]. There are backward connections going from the thalamus to the cortex, yet these are less data intensive. The cortex may provide the main categories and memory storage while the thalamus may signal to what extent the current stimuli are related with goals and stored categories. Once again consciousness, intermediate cognition and goal generation appear to be closely linked. In the end, the autonomy and adaptability of mammal brains might be endorsed by this general neural architecture devising generality and goal generation [8][9].

C. Amygdala

The amygdala is that part of the brain dealing with the emotional states of the subject. It contains a huge number of nuclei and receives projection from the cortex areas and return them. Adolphs and Spezio show that the amygdala plays an important role in guiding social behaviours on the basis of previously developed goals and the available sensory information. Mostly important they also hypothesize that the amygdala makes an attentional modulation in somatosensory cortices [1]. This hypothesis is consistent with the experimental evidence showing that the brain bootstraps the generation of new goals taking advantage of hardwire criteria likely located in the amygdala [3].

D. Brain inspired model

The brain areas presented above need to be integrated in an unique computational framework, namely a neural network. Moreover, the interaction among these areas must be mathematically specified. Figure 1 shows how the areas can be coupled with the biological counterpart. The cortex is composed by several *Category modules* where each of them represents a specific subcortical area, processing the same kind of incoming sensory information. The thamalus is divided in different modules, called *Ontogenetic modules*, representing the developed goals through the interaction with the environment. It is worth noting that each Ontogenetic module is closely coupled with a single category module, providing a relevance measure of the incoming sensory data with respect to developed goals. The amygdala is modelled as a single global module, called *Global Phylogenetic module*, representing the innate criteria. It provides a relevance measure with respect to some specific patterns in the sensory space (e.g. the interest for coloured object).

III. GOAL-GENERATING ARCHITECTURE

The *Goal-generating* architecture represents the cognitive middleware and it is implemented by a network of *Intentional modules* (IMs) and by a single *Global Phylogenetic module* (GPM). The IM groups an Ontogenetic module (OM) and the corresponding Category module (CM), modelling the interaction between the thalamus an a single sub-area of the cortex whereas the GPM models the amygdala, providing pre defined innate criteria. The GPM projects to each of the IM in the network, providing a global level of interest

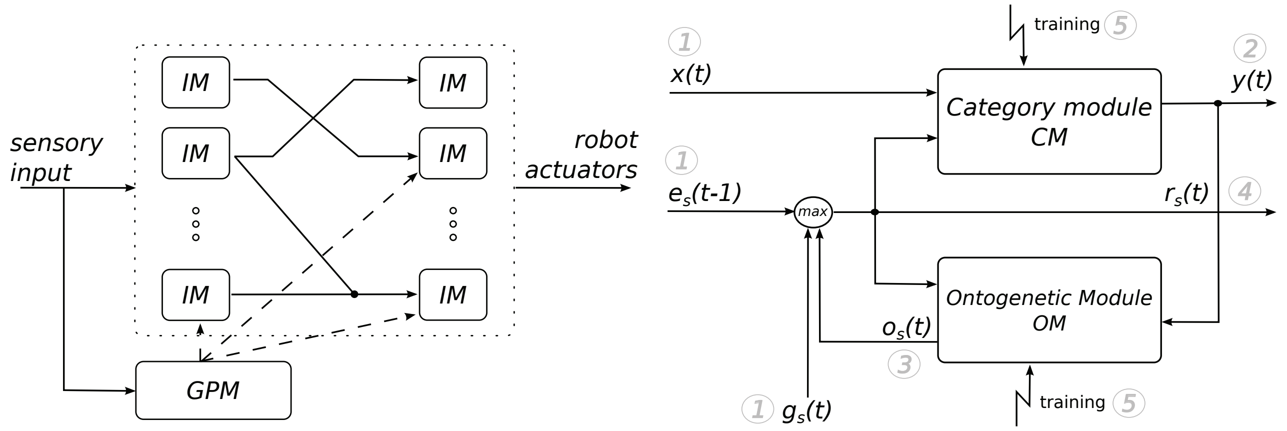

Fig. 2. (Left pane) The overall Goal-generating Architecture (IA). The network nodes are a set of Intentional Modules (IM) whereas the Global Phylogenetic Module (GPM) provides the relevance of the incoming sensory information. It receives in input a set of sensory information and produce as output the robot actuation. (Right pane) It shows the sub modules of the IM. It is composed by an Ontogenetic module (OM) which stores the developed goals and by a Category Module (CM) that produces an internal representation of the incoming sensory data. The numbers represent the discrete steps of the algorithm.

with respect to the filtered sensory information. The Goal-generating architecture receives as input the sensory data coming from the available sensors of the robot whereas the output is projected directly to the actuators (see Figure 2, left pane).

A. *Intentional Module*

The IM is the core module of the network and represents a part of the whole interaction among the cortex, the thalamus, and the amygdala (see Figure 1). The IM receives in input the sensory data $x(t)$ (eventually preprocessed), the global signal of relevance $g_s(t)$ from the GPM, representing the relevance of the current stimuli with respect to the innate criteria, and an external signal of relevance $e_s(t-1)$ coming from another IM (if any). It returns as output the category signal $y(t)$ representing the similarities of the current sensory data with previously stored information and the relevant signal $r_s(t)$. The signal $r_s(t)$ is important for several reasons: first it represent the relevance of the incoming sensory information with respect to the innate criteria and developed new goals and second it is fundamental during the learning phase for the develop of new goals (see Figure 2, right pane).

The IM is composed by due sub modules:

- the Category Module (CM)
- the Ontogentic Module (OM)

The CM models a cortical sub area and it is involved in the categorization of the sensory information whereas the OM models a thalamus sub area, computing the relevance of the incoming sensory information with respect to previously developed goals.

B. *Category module*

The CM addresses the representation of the sensory input generated either by a single source or a combination of them. The CM must be able to represent a mixature of sensory information regardless their unity of measure. We have implemented an ad-hoc clustering algorithm with a proper measure of distance based on the statistical correlation index. Figure 3 (left pane) show a sketchy outline of the CM. It receives in input the sensor (numerical) values and computes the distances of the current sensor input with respect to the developed clusters. Defining the input vector $x(t) \in R^M$ and the cluster centers $C_1, C_2, \ldots, C_N$ in the M-dimensional space, the CM computes the output vector $y(t) \in R^N$ as following:

$$y_i = \rho(x, C_i) \tag{1}$$

where $y = [y_1, y_2, \ldots, y_N]$ and $\rho(x, C_i)$ is the positive correlation index. The element y_i shows the similarity of the input stimuli with respect to the center i. For our purposes, we consider the vector y as a good candidate to represent the incoming sensory data. One of the main feature of the CM is to dynamically create clusters in the input space, depending on the relevant signal computed internally by the IM. It means that if the current sensory input is associated to a high relevant signal and it becomes a candidate to represent a cluster. To determine if it is necessary to generate a new cluster, the CM considers if the relevant signal $r_s(t)$ is greater than an imprinting threashold $t_{imprint}$ and if the minimum distance among the input vector $x(t)$ and the centers C_i is greater than a distance threashold t_{dist}. If so, a new cluster with a center equal to the current value $x(t)$ is created. Technically speaking, the output dimension N is modified according to the updated number of clusters.

C. *Ontogenetic module*

The OM represents the associative memory of the IM. It receives as input the the category output $y(t)$ and produce as output an ontogenetic relevant signal $o_s(t) \in R$, representing the relevance of the current input with respect to the developed goals. During the training phase, it uses the relevant signal $r_s(t)$ to eventually update the goals. To model the new developed goals, the OM uses a set of internal variables, called *gates*, representing the weights of neurons with one input and one output (as shown in Figure 3, right pane). The gates are grouped in a vector state $g = [g_1, g_2, \ldots, g_N]$

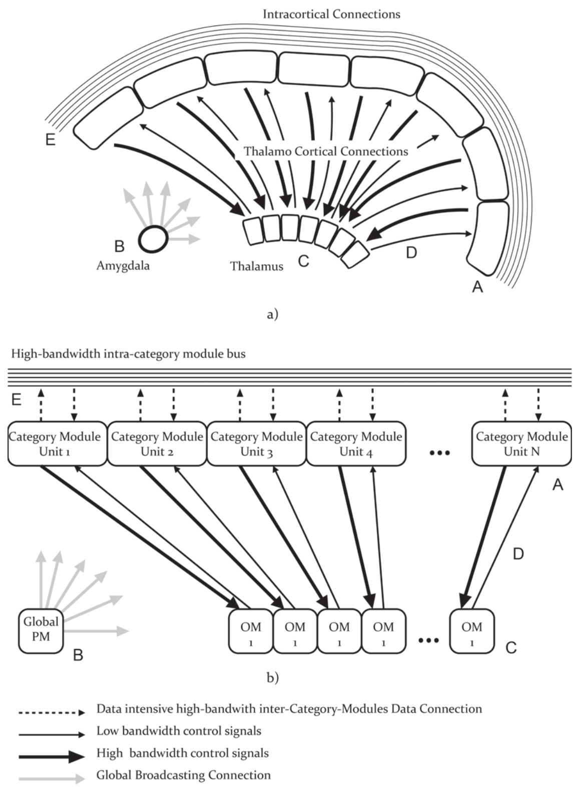

Fig. 1. A comparison between a (very) sketchy outline of the thalamo-cortical system as commonly described [11] and the suggested architecture. It is worth to emphasize the similarity between various key structures: A. Category Modules vs cortical areas, B. Global phylogenetic module vs. amygdala, C. Ontogenetic modules vs. thalamus, D. Control signals (high and low bandwidth), E. high bandwidth data bus vs intracortical connections.

where N is the exactly the y dimension. The $o_s(t)$ signal is computed as following:

$$o_s = \max_i (y_i \, g_i) \tag{2}$$

where y_i is the element value of the category output vector and g_i is the value of the associated gate. The signal $o_s(t)$ points out if the similarity of the clusters in the CM is relevant for the developed new goals. On the other hand, the gates g_i represents which cluster in the past was frequently associated with an higher relevant signal. To take into account the plasticity of the gates during the training phase, we implement a stable version of the Hebbian rule (Oja's version), in the following way:

$$\begin{aligned} g_i(t) &= g_i(t-1) + \Delta g_i(t) \\ \Delta g_i(t) &= \eta \left[r_s(t) y_i(t) - g_i(t-1) y_i(t)^2 \right] \end{aligned} \tag{3}$$

where η is the learning rate, $r_s(t)$ is the relevant signal that acts as output of the neurons i and $y_i(t)$ is the input of the gate. Moreover the OM has a saturation mechanism to take into account long term effects:

$$g_i > th_g \Rightarrow g_i = 1 \tag{4}$$

where th_g is a value representing a long term effect of correlation between $x(t)$ and $r_s(t)$. The gates encode the new developed goals, making an association between the clusters in the CM and the past history.

D. Signal propagation

The signal propagation through the internal sub modules of the IM can be summarized in the following steps (see Figure 2, right pane):

1) The sensory input $x(t)$, the global relevant signal $g_s(t)$, and the external relevant signal $e_s(t-1)$ of the previous time step are received;
2) the CM computes its output signal $y(t)$ which encodes a proper representation of the actual sensory signal x(t);
3) the OM computes the relevance signal $o_s(t)$, knowing the actual input representation $y(t)$ and having an internal representation of the developed goals in the input space;
4) given $g_s(t)$, $o_s(t)$, and $e_s(t-1)$, the IM computes its relevant signal $r_s(t)$ that represents the relevance of the sensory input at time t;
5) *training phase*: the internally generated relevant signal $r_s(t)$ (in the previous step of this algorithm) drives the developing of new categories in the CM and the updating of the gates in the OM.

E. Global Phylogenetic Module

The GPM contains the innate criteria driving the bootstrap of the overall architecture. It receives in input the whole sensory information coming from the robot and computes the $g_s(t)$ signal, eventually forecasted to every IMs in the network. It is worth noting that the IMs typically receive in input only a small subset of the available sensory data, according to cortical sub areas that deal only with specific sensory input, e.g. primary visual cortex.

The GPM is composed by several *phylogenetic functions* that interact to compute the global signal $g_s(t)$. Phylogenetic functions serve as a repository of ready-to-use criteria. For instance, it is well known that in humans the face recognition is triggered by some innate criteria that in our case are modelled as phylogenetic functions. Mathematically speaking, the GPM can be modelled as follow:

$$g_s(t) = \max_i f_i^p(S_x(t)) \tag{5}$$

where f_i^p is the *i-th* phylogenetic function and $S_x(t)$ is a subspace of the incoming stimuli $x(t)$.

IV. Discussion

The proposed architecture is the result of some theoretical improvements described in Section II and III. Before to extend the experimental setup, we would validate previous results obtained in [9]. Those results are obtained with a preliminary version of the cognitive architecture that is presented here.

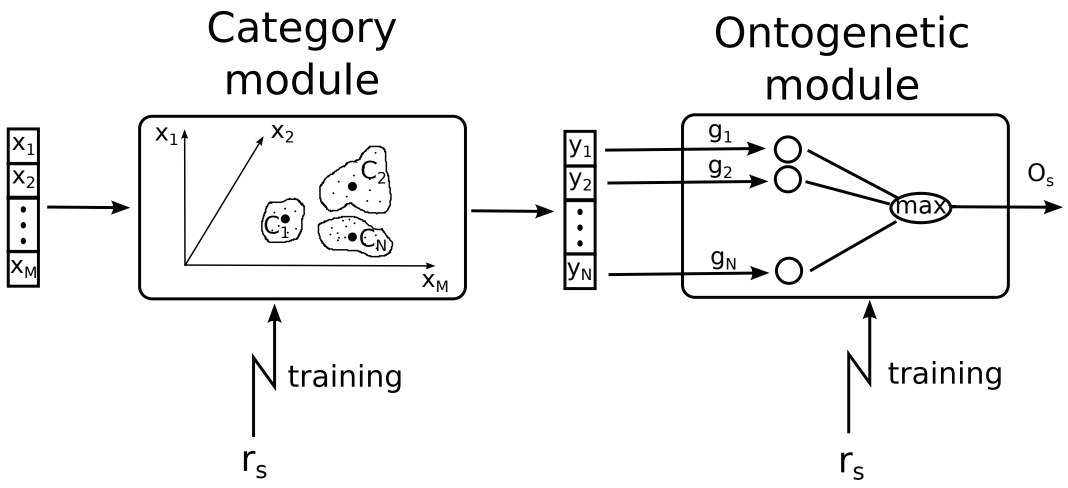

Fig. 3. (Left pane) The category module (Right pane) The ontogenetic module

A. Setup

Here a robotic setup, based on a standard commercial robot NAO (Aldebaran Robotics) with 21 d.o.f. is taken into consideration. The aim of the setup is to check whether the above mentioned recipe for generating a cognitive architecture is efficacious in a real environment. Efficiency is not an issue here. The goal is to check whether a whole architecture may be generated by unconstrained interaction with the environment, without being explicitly defined at design time. Potentially the goal-generating architecture can control a generic robot regardless of its structure.

In this specific experiments the overall architecture is composed by a single Intentional Module and by the Global Phylogenetic Module. The GPM contains a single phylogenetic function and, given the Equation 5, the global relevant signal is equal to the function value $f_i^p(S_x(t))$, where $S_x(t)$ is the incoming RBG frame from the frontal camera and $f_i^p(\cdot)$ is the normalized saturation value of the incoming frame. More the image contains coloured object and more the saturation value will be higher. The input signal $x(t)$, that is received by the IM, is the log-polar transformation of the incoming image frame.

To replicate the experiment, we used only the pan and tilting commands of the NAO's head while it is sitting in a stable position. The angle change is given by a Gaussian distribution, centered at the actual value, with a spread that depends on the relevant signal: the higher the signal, the smaller the movements. In this way the experimenter may have a visual feedback of the relevance of a particular image for the robot.

At the beginning, the architecture is empty and its modules are empty. Then the system is switched on. The IM begins to receive new data and, accordingly to the hard-wired bootstrapping criterion, begins to burn the elementary units thereby beginning to get coupled with the particular environment.

B. Goal-generating behaviour

The aim of the experiment is to verify if the robot is able to generate a new goal starting from an hard-wired innate criterion. At the beginning, the robot generates a positive phylogenetic signal when it sees a coloured object regardless its shape and what we want to obtain, after the interaction with a structured environment, is a robot that produces a relevant signal also for a specific object shape regardless its color. The experiment procedure is divided in three steps:

1) a set of images containing geometrical black objects with different shapes are shown in front of the robot
2) we present in front of the robot a specific image containing a brightly colored object with a very specific shape (e.g. a star). The image is maintained empirically for a certain amount of time.
3) in the last stage, the same set of image presented in the first step are re-proposed. They will contain at least a black object with the same shape of the object proposed at the second step.

What we expect to observe is an emergent behaviour where the cognitive architecture will produce a positive relevant signal also for a particular object shape (e.g. a star) and not just only for coloured object. Therefore, the GPM generates a positive signal for coloured object and the Category module implements the clustering algorithm to group the sensory vectors in clusters where the relevant signal drives directly the head joints.

Following the experimental protocol, during the first phase the robot explores with long steps the surrounding environment without showing a particular interest for a region of the space.

During the second step we present to the robot an image with a brightly coloured object shaped of a star. Immediately the phylogenetic signal becomes high. However, only after a phase of associative learning the ontogenetic signal becomes higher, too. Moreover, the head movements become very

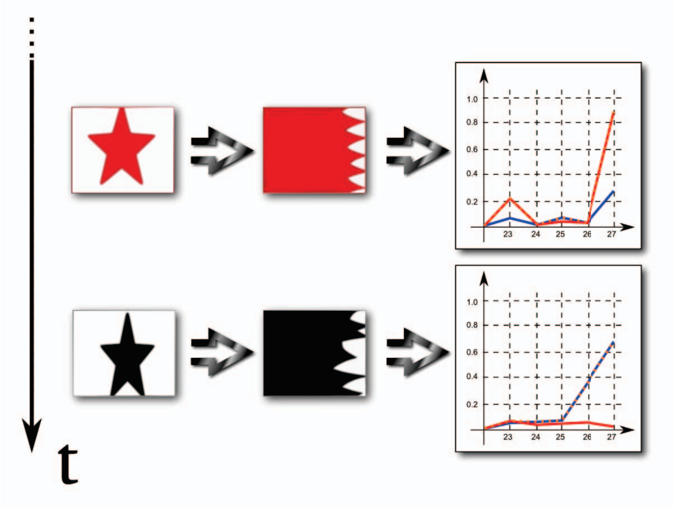

Fig. 4. Performed experiment. The red line represent the global relevant signal whereas the blue line represent the ontogenetic signal. The IM receives in input the log-polar image of the scene. The learning is performed in the log-polar space. After a while, the cognitive architecture is able to produce a signal of relevance also for the star-shaped objects.

small thereby keeping the gaze close to the center of the coloured object. Finally, in the third phase, we presented again the original black objects. From the experience of the first step we should notice that, in general, the robot does not focus on black objects. However, eventually the agent shows interest for star-shaped black objects. In this case the high value of the relevant signal is due to the ontogenetic module. In fact, in the previous step, the ontogenetic module has associated, through the category module, the shape of the object with the phylogenetic signal. At the same time, we do not observe the same behaviour with black objects of different shape (see Figure 4).

V. Conclusion

In this paper we have proposed a cognitive architecture aiming to generate new goals starting from a single bootstrap criterion, only with the interaction with the environment. Notwithstanding the current shortcomings, the architecture is able to generate a new goals starting from an hard-wired criterion. This paper introduce some improvements with respect to a previously goal-generating architecture. First, the Hebbian learning is introduced to learn new goals and second a Global Phylogenetic Module (GPM), modelling the amygdala, is proposed as a global system that forecasts to each node of the network the relevance of the current input stimuli, following some recent evidences in the brain. In the future we plan some improvements. First we investigate different way of input representation, such as radial basis functions, and Independent Component Analysis. Second we investigate the generation of new Intentional Module and the connections plasticity among different Intentional Modules in the network. Finally, we investigate the capability of the cognitive architecture to infer movements that maximize its goals.

VI. Acknowledgments

This work has been possible thanks to the support by the Italian-Korea bilateral project between ICT-CNR and KAIST (2010-2012).

References

[1] R. Adolphs and M. Spezio. Role of the amygdala in processing visual social stimuli. *Progress in Brain Research*, 156:363–378, 2006.

[2] G. Dileep. *How the brain might work: A hierarchical and temporal model for learning and recognition*. PhD thesis, Stanford University, 2008.

[3] S. Duncan and L. F. Barrett. The role of the amygdala in visual awareness. *Trends in Cognitive Sciences*, 11(5):190–192, 2007.

[4] J. Hawking and S. Blakeslee. *On Intelligence*. Times Book, New York, 2004.

[5] A. Hyvrinen and P. O. Hoyer. A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Research*, 41:2413 2423, 2001.

[6] A. Hyvrinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, 13:411–430, 2000.

[7] R. S. Jackendoff and F. Lerdahl. The capacity for music: what is it, and what' s special about it? *Cognition*, 100:33–72, 2006.

[8] R. Manzotti. *A Process-oriented Framework for Goals and Motivations in Biological and Artificial Agents*. Ontos-Verlag, Frankfurt, 2010.

[9] R. Manzotti and V. Tagliasco. From behavior-based robots to motivation-based robots. *Robotics and Autonomous Systems*, 51:175–190, 2005.

[10] D. S. Modha and R. Singh. Network architecture of the long-distance pathways in the macaque brain. *Proceedings of the National Academy of Sciences*, 107(30):13485–13490, 2010.

[11] R. Nieuwenhuys, J. Voogd, and C. van Huijzen. *The Human Central Nervous System: A Synopsis and Atlas*. Steinkopff, Amsterdam, 2007.

[12] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996.

[13] J. Sharma, A. Angelucci, and M. Sur. Induction of visual orientation modules in auditory cortex. *Nature*, 404:841–849, 2000.

[14] M. Sur, A. Angelucci, and J. Sharma. Rewiring cortex: the role of patterned activity in development and plasticity of neocortical circuits. *Journal of Neurobiology*, 41:33–43, 1999.

[15] M. Sur, P. E. Garraghty, and R. A. W. Experimentally induced visual projections into auditory thalamus and cortex. *Science*, 242:1437–1441, 1988.

[16] L. M. Ward. The thalamic dynamic core theory of conscious experience. *Consciousness and Cognition*, 20:464–486, 2011.