Advanced Robotics

ISSN: 0169-1864 (Print) 1568-5535 (Online) Journal homepage: http://www.tandfonline.com/loi/tadr20

From learning to new goal generation in a bioinspired robotic setup

Alessio Mauro Franchi, Flavio Mutti & Giuseppina Gini

ADVANCED ROBOTICS, 2016
http://dx.doi.org/10.1080/01691864.2016.1172732

FULL PAPER

From learning to new goal generation in a bioinspired robotic setup

Alessio Mauro Franchi, Flavio Mutti and Giuseppina Gini

Dipartimento di Elettronica, Informazione e Bioingegneria Politecnico Di Milano, Milan, Italy

ABSTRACT
In the field of cognitive bioinspired robotics, we focus on autonomous development, and propose a possible model to explain how humans generate and pursue new goals that are not strictly dictated by survival. Autonomous lifelong learning is an important ability for robots to make them able to acquire new skills, and autonomous goal generation is a basic mechanism for that. The Intentional Distributed Robotic Architecture (IDRA) here presented intends to allow the autonomous development of new goals in situated agents starting from some simple hard-coded instincts. It addresses this capability through an imitation of the neural plasticity, the property of the cerebral cortex supporting learning. Three main brain areas are involved in goal generation, cerebral cortex, thalamus, and amygdala; these are mimicked at a functional level by the modules of our computational model, namely Deliberative, Working-Memory, Goal-Generator, and Instincts Modules, all connected in a network. IDRA has been designed to be robot independent; we have used it in simulation and on the real Aldebaran NAO humanoid robot. The reported experiments explore how basic capabilities, as active sensing, are obtained by the architecture.

ARTICLE HISTORY
Received 14 July 2015
Revised 15 January 2016
Accepted 24 March 2016

KEYWORDS
Goal generation; active perception; learning; cognitive development; intrinsic motivation

1. Introduction

Human beings are able to develop several mental abilities during their life. This process is called cognitive development and refers to those mechanisms as perception, thinking, and understanding of the world; it is a mixture of both genetic and learned factors.[1] How humans are able to develop these capabilities during their existence is not completely understood, but the autonomous generation of new goals and behaviors upon simple innate criteria appears as a crucial factor in this evolution, to allow individuals to adapt their behavior to the various situations they face everyday.

In order to build artificial agents capable of interacting in an effective way with humans and to be really integrated in our life, robotics might take inspiration from those processes that allow our brain to show this cognitive development. The here presented work intends to contribute to the achievement of this objective: its purpose is to create a bioinspired agent, based on a simplified model of the brain functionalities that characterizes the autonomous development of new goals, and able to learn new behaviors that could attain these goals.

The problem of robot's adaptation has been addressed in several ways in literature. A first attempt is behavior-based robotics, which states that the agent should adapt its behavior to the environment in order to accomplish its goals, but these internal objectives are only the hard-coded ones and no others may be learned during the agent's life.[2] On the contrary, developmental robotics tries to imitate the emergence of cognition in natural and artificial systems.[3] Developmental robotics aims at the cognitive development of the agent making it able to adapt to the environment and to autonomously develop new motivations that were not present at design time.

Autonomous development and lifelong open-ended learning are observed in mammals, and especially in humans, which engage in activities that are not directly aimed at survival.

Manzotti and Moderato [4] argues about a forthcoming science of intentional changes; it focuses on two paradigms that have been developed in AI and psychology: intrinsic motivation and Hierarchical Open-ended Architectures (HOA).[5,6] Interesting for our purposes is that it addresses the issue of learning new motivations. Intrinstic motivation and HOA are closely related to the debate among the evolutionary psychology model (that privileges innateness) versus the social science model, which endorses the open-ended capacity for change.

Intrinsic motivation aims to model doing something because it is inherently enjoyable.[7,8] Intrinsically motivated behaviors can result from innate rules such as maximizing novelty, or be the outcome of an open-ended architecture. Machine leaning algorithms (RL), as reinforcement learning, are able to incorporate some aspects

CONTACT Giuseppina Gini giuseppina.gini@polimi.it

of intrinsic motivations; they have been applied in robotics, in particular for the acquisition of motor skills, to increase the autonomy of artificial systems with motivations that can support lifelong autonomous learning. Bioinspired computational models are linking some neuroscientific discoveries to specific computational mechanisms.[7] Fiore et al. [9] proposed a computational model of the basal ganglia that uses a unique signal to obtain both intrinsic and extrinsic motivations, and tested it in manipulation tasks.

Considering its internal status and also the implicit motivation of fulfilling innate instincts, the robot should carry out the selection of the action; this kind of internal motivation has been widely covered in [10]. An interesting study tried to model both curiosity and manipulation in a robotic setup to drive the robot in understanding the dynamic of objects [11]; the explorative behavior shown by the robot is motivated by novel phenomena and is modeled as a probability system. Other computational mechanisms related to intrinstic motivation are being proposed in the field of active learning, in particular in relation to supervised learning systems.[12] Another different approach exploits the theory of Motivated RL, in conjunction with different models of attention and goal-oriented behaviors; [13] presents an architecture able to learn both an introspective policy for when an agent should create and delete goals, activate or suspend skills, and a multiple-option policy for mapping states to actions for those skills. An implementation for controlling non-player characters in computer games has demonstrated that the addition of introspection helps in focusing attention on more complex goals and in learning quickly, but no tests on robots have been performed.

Let us now go back to HOA. Many psychological approaches apply traditional learning paradigms (as the operant conditioning) to explain how an intentional cognitive agent may produce a new goal. Being motivated and developing new goals are different, but both involve learning. Here is how HOA plays a role: to take decisions an agent must be able to represent external stimuli and states, in an unlimited hierarchy, where new concepts and motivations could be developed. Openess is a key property for a cognitive architecture to generate new goals.

Our approach here is different in some ways from those previous works. We want to define a system using some kind of internal motivation and able to autonomously generate new goal or interests starting from a simple set of innate instincts and imitating the functionalities of the brain. Our theory is loosely inspired by the perceptual control theory (PCT), an early social science model for human behavior stating that 'it is perception which gives form to behavior'.[14] PCT experiments have shown that an organism controls neither its own behavior, nor external environmental variables, but rather its own perceptions of those variables. In our experiments, we indeed see a natural emergence of an active perception like behavior, which is known to be fundamental in robotics since the landmark paper by Bajcsy [15]; active sensing means that the perception process integrates purposeful actions and sensing so to improve data interpretation according to the task.[15]

With our research we specifically address an intermediate level of cognition aimed at allowing mammals and humans to be aware of the surrounding environment, and then to interact with it, even without very complex reasoning. We want to focus on how to make cognition emerge in agents and robots, not caring about how and where to drive them, a concept called 'artificial curiosity'; this terms implies that artificial agents should explore and discover how the external world works, interacting with it through the sensor-motor interface.[16] A first attempt of embodying the concept of curiosity in a humanoid robot showed that it can improve and accelerate the learning of its configuration space.[17]

We claim that the capability of goal generation is an essential precondition to enable robots to fit into humans everyday life; what is missing nowadays in robots is this skill of acting in a consistent manner with respect to changes in both the surrounding environment and their own structure; agents able to develop new goals emerging from manipulating the environment would be able to effectively interact with people.

Our work takes inspiration from nature, in particular from the high-level structure and communication flow of data in the human brain: three areas of the brain and their interconnections are involved in the cognitive development process: cortex, thalamus, and amygdala.[18] See Figure 1 for a schematic representation of the brain and its transposition to the proposed model. Our Intentional Distributed Robotic Architecture (IDRA) is an open network of elementary units, called deliberative modules (DM), which enables this active learning process. The network is open, and is composed by several layers connected both in feedforward and feedback mode.

Concepts developed in machine learning (ML) seem to find their correspondence in the brain, from structural organization to the neuronal level. The ML perspective deals with states and values, whereas the neuronal perspective works on finding and interpreting neuronal signals. Take the example of reinforcement learning (RL); the agent learns from the consequences of its actions, and it selects its actions on the basis of its past experiences as well as for trying new choices. For ML the reinforcement signal is a numerical reward, which encodes the success of an action's outcome. Looking for a biological mechanism for reward, we expect that the reward information be processed by specific neurons in specific brain areas. Some of the individuate neurons are indeed in the cortex and in the amygdala; we are not modeling their activation, but instead only their connections.

Beside the DMs a single instincts module (IM) contains the hard-coded objectives, i.e. the innate instincts, exactly as in the amygdala. The IM spreads all over the network a signal which reflects the emotional state of the agent with respect to genetic criteria: the more the current state of the robot meets its own innate goals, the higher is the signal. Each DM is in turn composed of two sub-modules: working-memory (WM) and goal-generator (GM) module. WM acts as the cerebral cortex and returns the cortex activation in response to the actual sensorial input. The GM finally performs the generation of new goals; it receives the values from WM, and through Hebbian learning it lets new objectives emerge; this module generates a scalar signal in the [0 1] range indicating how much the current state meets these new goals.

To execute the goals, the motor system (MS) may be activated. At each cycle, the network of DMs outputs a couple composed by a vector and a scalar; it is sent to the MS, which should generate movements consistent with the goals of the agent. A movement is a composition of a series of elementary signals, called motor primitives (MPs), which represent muscular activations over time; this muscular synergy leads to the execution of complex movements.[19,20]

We already illustrated the potentiality of IDRA in classification tasks, both on visual and audio signals [21]; in the present paper instead we devise two experiments to verify its goal generation capability as well as its ability in learning how to adapt its behavior to new objectives. In the first experiment, the robot learns how to distinguish a particular shape starting from an innate instinct to be attracted by highly saturated figures; in the second experiment, we focus on the active-sensing tasks: the NAO robot has to autonomously explore and learn new movements of its arm in order to satisfy its instincts.

The main contributions of this work are:

- the design and full implementation of a cognitive architecture based on an amygdala-thalamo-cortical model, as proposed in [22];
- the integration of an Intentional Architecture and a MS;
- the exploration of active-sensing paradigm for learning;
- the validation of the architecture by testing it on robotics tasks.

Our proposal is a step in the direction of trying to integrate psychological and computational methods, and to investigate how it can impact the design of autonomous robots.

2. A computational model of the Amygdala-Thalamo-Cortical structure

The IDRA architecture takes inspiration from the amygdala-thalamo-cortical circuit in the brain at its functional level. Several studies in literature have shown the importance of these brain areas in cognitive development, which is a network phenomenon, and does not exist in synapses or single neurons.[24] In the following, we sketch out the most significant functionalities of those three involved areas and their interconnections (see Figure 1).

The cortex is the external part of the brain; it is divided into two hemispheres and receives signals from the sensory organs. There is a specific section connected to each sensory input, e.g. the visual cortex or the auditory cortex, however different cortical areas can properly react to different stimuli sources,[6,25] since the statistics of the incoming signal is the key for the cortex to adapt to new inputs.[26] The whole cortex is composed of the same kind of cells and therefore each cortical area has the ability to virtually implement any computational skill.[27,28] For imitating this interesting capability of the cortex we sought for a general approach for representing stimuli that valid candidate for this high-level representation.[29] ICA is a powerful way to analyze multi-variate data, as sensorial inputs are, and is able to learn a generic decomposition of sources based on their statistical properties, exactly as it happens in human brain; this data-driven analysis can be used to describe data at an higher level.[30]

The thalamus is located deep in the brain; it plays a central role for mammals in the development of new motivations, as well in the choice of what goal to pursue. The thalamus is a ‘central, convergent, compact miniature map of the cortex’ [31]; its structure is partitioned into segments, each one in synchronized projection to a specific sub-area of the cortex. Since the latter carries on most of the data processing, storing and distribution, the thalamus must provide to the corresponding cortical area which goal have to be pursued.[32]

The amygdala is located within the medial temporal lobes; it is heavily connected to the cortical areas and is involved in the generation of somatosensory response on the basis of innate goals.[33] Experimental evidences

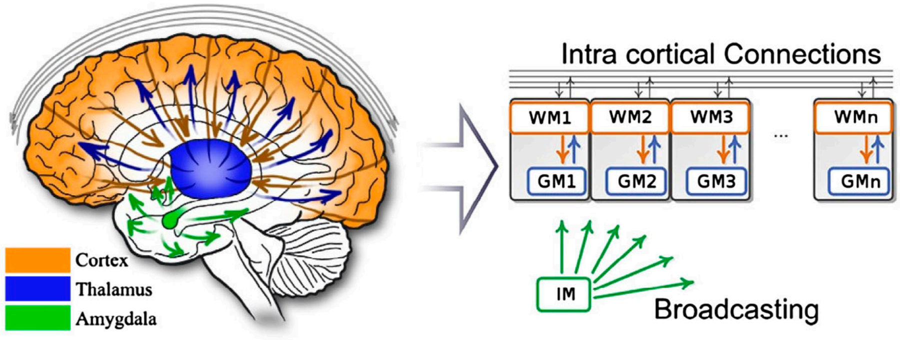


Figure 1. A simplified structure of the human brain [23] and its transposition in our architecture.

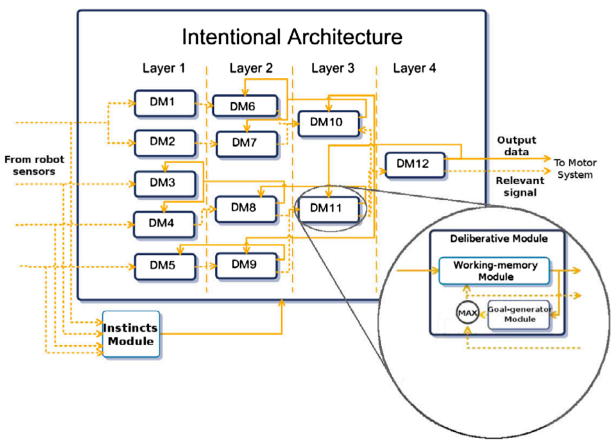


Figure 2. Several DMs are interconnected in a layered network; each DM is composed by two sub-modules: WM and GM; a IM feed each DM with an emotional signal.

show that the brain bootstraps the generation of new goals taking advantage of the innate criteria located in the amygdala.[34]

Asides from these three areas, this work approaches basic concepts toward the generation of motion for enabling the so called sensorimotor circuit. That includes the cerebellum [35] and the spinal cord.[36] The motor primitives model is widely referred to as the most valid model for explaining motor learning; it refers to the generation of complex trajectory by combining a limited set of waveform modules.[37]

3. The IDRA architecture

The simplified functional model presented in the previous section has been transposed in a ‘goal-generating’ architecture, that is the IDRA architecture.[22] It is essentially a network of DMs plus a single IM, which simulates connections and interactions between the cerebral cortex, the thalamus and the amygdala.[38] Figure 2 shows a schema of the whole architecture.

Each DM is in turn composed by one WM and one GM, modeling the interaction between the thalamus and a single sub-area of the cortex. All the DMs may be linked in any way, while the IM broadcasts its signal to all the DMs, without receiving data back. This network has one or more input coming from all the available sensors, possibly filtered to extract meaningful features, and two outputs: one vector representing the level of activation of the WM generated by the current sensory input, and a scalar signal stating how much the actual input satisfies both instincts and generated goals.

The architecture, thanks to the GM (that emulate the functionalities of the thalamus), is able to autonomously generate new goals upon some simple innate criteria; these lasts are called instincts and are implemented as hard-coded functions in the IM (the amygdala of the system), whose output signal tells the agent its emotional state with respect to instincts.

The WM acts as the cerebral cortex; it receives input from sensors or from other DMs and performs unsupervised categorization. The input is elaborated twice: first with independent component analysis (ICA),[29] then with a clustering algorithm. ICA allows to abstract from the type of the incoming stimuli. During a training stage several input data examples are presented to the robot so that it learns a decomposition of that particular type of input (i.e. several independent components are extracted from data); at runtime each new input is projected in the n-dimensional space of the previously created independent components to both reduce the dimension of the problem and build a general representation of the incoming data. Equation (1) formalizes this projection:

$$\overline{W} = \overline{IC} \times \overline{I} \tag{1}$$

where $\overline{W}$ is a set of weights representing the projection of the raw vector input $\overline{I}$, and $\overline{IC}$ is the matrix of independent components. The next step, i.e. clustering of input data, is performed on this set of weights $\overline{W}$ instead of on raw data; the result is a code of the input. During this process, each vector $\overline{W}$ is assigned to an existing cluster if its distance from that is below a previously-set threshold, otherwise a new cluster is created, and the newly acquired vector is its centroids. A maximum number of cluster for each WM is set a priori to 128 in out tests. The output of the WM is a vector containing the activations of each stored clusters, which is inversely proportional to the distances of the new input data from the each centroids as in Equation (2):

$$y_i = \rho(x, C_i) \tag{2}$$

where y_i is the distance of the input x from the center C_i of cluster i. The generation of new categories depends also on the internal relevant signal; only relevant inputs are categorized, so that the module learns only meaningful information. With the term relevant we are referring to input data that excite the most either the IM or the GM, meaning, respectively, that they satisfy hard-coded criteria or generated goals. Only input that generate an internal signal greater than a fixed threshold are considered relevant; the threshold value is set empirically so to have a good trade-off between speed and quality of learning.

Finally, the GM is closely connected to the WM and it uses the activations of categories computed by the WM and an Hebbian learning function to develop new goals. It returns a signal os stating how much these new goals are satisfied, calculated as in Equation (3):

$$os = \max_i (y_i w_i) \tag{3}$$

where y_i is the activation of the ith category in the WM and w_i is its corresponding gating weight ($w_i \in [0 \dots 1]$).

These weights are updated at each iteration using an Hebbian learning function (Equation (4)):

$$w_i = w_i + \eta(h_s y_i - (w_i y_i^2)) \tag{4}$$

where η is the learning rate and h_s is a control hebbian signal we fixed to one in our experiments. We expect that the os signal after random small variations will increase according to the action taken by the robot. High values of os denote the emergence of new goal.

The output of the network is the vector produced by WM and the scalar produced by GM; they are sent to the MS. Here we propose a solution based on the 'State-Action' table for movement evaluation and on the concept of Motor Primitives for the composition of complex movements.

Several scientific evidences have led to the idea that movements are composed of elementary building blocks, called motor primitives and voluntary actions are generated by bonding them to each other either simultaneously or serially in time (Figure 3).[19,20] Following this idea, we use motor primitives to create complex muscular activations. A motor primitive could be seen as the activation of a muscle during time: the higher the value of the primitive, the stronger the muscle activation. We implement primitives as Gaussian functions delayed in time (Equation (5)):

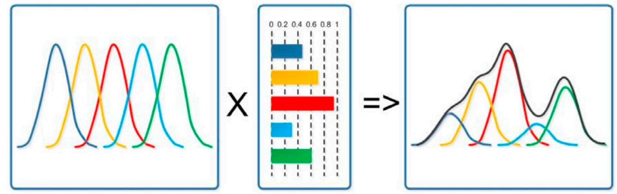


Figure 3. Motor primitives composition: each Gaussian function is a motor primitive; the weighted sum of all these primitives is the final movement; learning a movements is thus learning an optima set of weights.

$$p = e^{\frac{-(x-c)^2}{2\sigma^2}} \tag{5}$$

where c is the center of motor primitive p and σ is its amplitude. Then a movement m may be described as in Equation (6), where w_i is the ith weight:

$$m = \sum_i w_i p_i \tag{6}$$

The first step for movement generation is a new clustering of the input state via a K-means algorithm, using the clusters defined during a training phase. Given the current state $\overline{Y}$, the 'State-Action' table is able to learn and select the best movement to be performed; this table associates a state $\overline{Y}$ and a movement (in term of weights w_i of motor primitives) to the relevant signal coming from the architecture (see Figure 4). When the system is in a certain state and performs a particular movement, the output relevant signal is stored in the table; this signal is treated as a simple reward for an action. There are three different possibilities when a new input is read:

(1) if there is a movement associated with a relevant signal above a defined threshold, that movement is selected (memory skill);
(2) if there is a movement not yet performed, that movement is selected (bootstrapping);
(3) if all movements have already been performed at least once and no one is associated with an high relevant signal a new random movement is added to the list (explorative behavior).

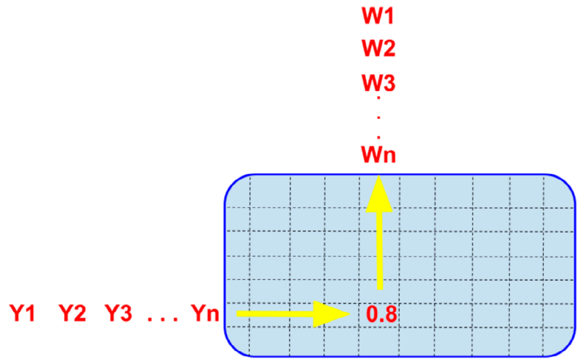


Figure 4. The State-Action Table; $Y_1, \cdots, Y_n$ represents the actual state, $W_{m1}, \ldots, W_{mn}$ are the weights for movement generation; each entry is the reward of the action, i.e. the Relevant Signal.

The proposed model has been implemented in the 'IDRA Software'; a brief outline of its structure is presented in the Appendix 1. The software is open and may be freely examined and used; for any question please refer to the first author of this paper.

4. Experimental results

This section describes the experimental phase; the purpose of these tests is to check whether the IDRA architecture here described is able to generate new goals and to learn consistent behaviors (i.e. movements), starting from simple hard-coded criteria. The IDRA architecture may potentially be used for controlling any kind of robot, regardless its type, its available sensors and actuators; we select the NAO, a humanoid robot by Aldebaran Robotics, with 25 degrees of freedom (DOF) and many sensors, including two cameras, four microphones and two sonars.

Two experimental setups have been prepared: the first focuses on sensing and deals with shapes and colors, as reported in [18,22], the second explores the creation of new movements in response to a specific sensorial input. Tests have been repeated several times each with different parameters: initial states and robot initial head and arm position to avoid bias. Statistics about the convergence of the learning phase have been collected; since it always appears at about 1000 internal cycles of the sensing-acting loop, we do not list all these data.

4.1. Learning from sensing

For the first experiment IDRA has a single DM and two filters in the input stage, the first computing the log-polar transform[1] of the visual input from NAO's top camera, the second extracting the overall saturation of the same image. The only innate instinct hard-coded in the amygdala is the attraction for colors, i.e. the robot is excited when it clearly sees high saturated figures or objects. As DOF we select the head movements, namely HeadPitch and HeadYaw. A simple attention mechanism is active: the greater is the interest of the robot for what it is looking at, the smaller is the angular head rotation in both the directions. The visual input for this experiment is made of two boards alternatively put in front of the top Nao's camera: the first board presents a series of black shapes, among which there is a black star; the second board presents only stars filled with three different highly saturated colors.

This setup allows the robot to explore the environment moving its head, looking for something interesting, either from innate or learned point of view: an active sensing behavior is created.

The test consists in three consecutive phases. Initially we present to the NAO the first board with only black figures; the interest of the robot is equally spread all over the board, not showing any particular interest for a specific shape. Both internal signals are very low and no interesting input data are clustered in the WM. Then the second board is shown; the robot, driven by its innate attraction for colors, focuses on the three star-shaped figures; the IM signal is now high and 128 categories (the maximum allowed) are created in the WM after 1000 cycles; moreover the GM signal grows indicating that new objective are starting to be generated. Lastly, the board is switched again to the first one and the interest of the robot is now mostly focused on the star-shaped figure, even if it is totally black. Figure 5 shows the result; in the top half blue dots show where the robot is looking at with respect to the board; in the bottom half, for each stage, the red signal is the output of the IM and tells us how much instincts are satisfied, while blue signal comes from the GM and describes when new goals are generated. In the third phase, we observe that the red and blue signals are, respectively, low and high, confirming that the robot is acting only according to new goals, specifically an interest in the shapes of the figures that adds to the previous innate interest for colors. The figure shows the two signals from 0 to 2000 cycles.

4.2. Action for sensing

For the second experiment, we add the MS since we want the robot to make more complex movements; yaw and pitch of the head are now fixed and we consider the movement of the four joints of the right arm, namely RShoulderPitch, RShoulderRoll, RElbowRoll, RElbowYaw. The attraction for coloured objects is the innate instinct in the IM. For moving the arm we define five motor primitives as in Equation (5), setting $\rho = 6.7$ (empirically a priori

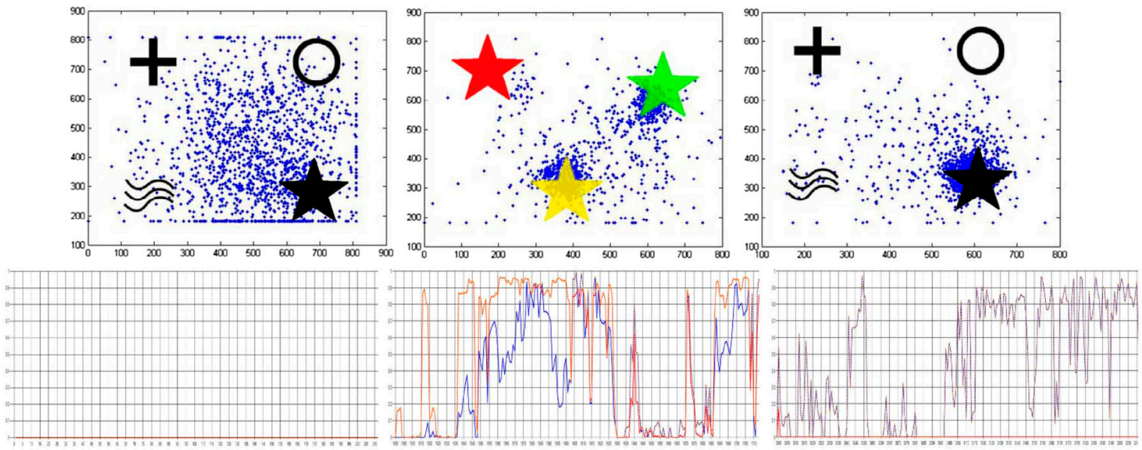


Figure 5. From left to right the results of the three phases of the first experiment; blue dots represents where the robot was pointing at, red and blue signals are, respectively, from the IM and the GM.

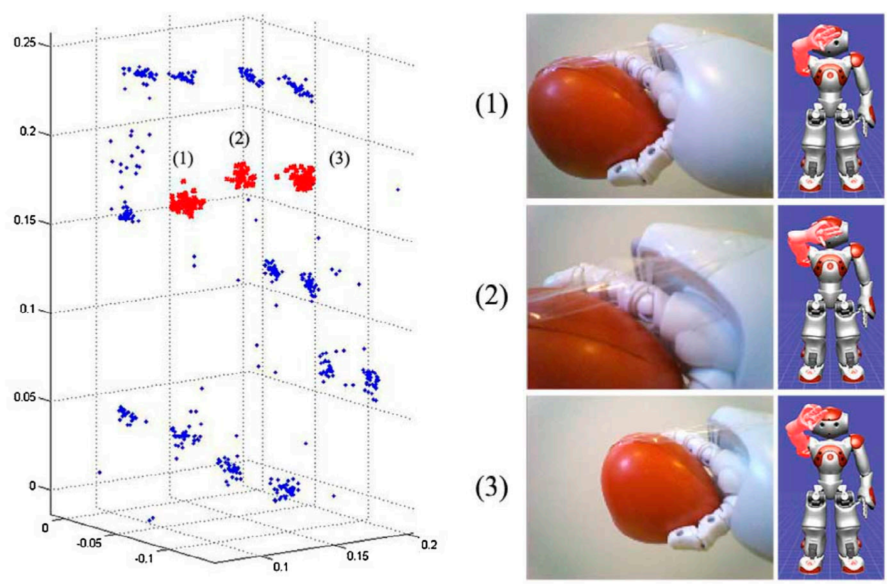


Figure 6. In the three-dimensional (3D) graph blue dots represent the positions of the right hand at each cycle, while in red are highlight those corresponding to good positions (as shown in the right part of the image); the higher concentration of dots in these three areas demonstrates that the robot has learnt various movements to accomplish its goal.

evaluated) and a mean value computed in order to equally distribute these functions on a scale from 0 to 100.

The log-polar and a saturation filter as in the first experiment, plus a third filter for proprioceptive data, called 'rightArmPosition', to retrieve joints values are used. Two layers compose the IDRA network: in the first there are two DMs, one for processing log-polar images, the other for receiving proprioceptive data; both send their output to the third DM in the second layer and its output is sent to the MS.

From this experiment, we expect the robot to start exploring its motion skills for fulfilling its goals randomly composing movements of its arm, and eventually learning how to perform appropriate complex movements in response to sensing. The attention is now mainly toward the robot exploiting this active sensing loop.

We set a simple objective: the robot has to look at the red toy it holds in its right hand; we do not hard-code any movement for the right arm, but its innate interest for colored objects should drive him in learning the right movement composition to get the red heart near the camera.

Each trial of the experiment starts with the right arm of the robot in a random position; the State-Action table is initially empty and movements are consequently chosen and executed randomly for bootstrapping the system. For each trial the system records in the table the associated *os* signal; after several movements the table starts filling up and several categories (about 100) are collected. The system may now choose movements coherently with the *os* signal received from the architecture; in our trials this bootstrap process took a few tens iterations, a variable value due to initial random action selection.

Going on with repeating this schema, the table is filled and so best movements start to be frequently repeated; the more a movement is performed, the more it is reinforced by the architecture.

Figure 6 shows the positions of the end effector of the robot in a 3D space during a thousand iterations; the frame of reference is centered in the body of the robot. We can see the emergence of several clusters; the red ones, associated with the highest relevant signal, have a visual 3D representation of the NAO. We can notice that for these configurations the ball is roughly at the center of the visual field of the robot, meaning that the movement was appropriate. The highest concentration of points in these three clusters means that the robot has autonomously learnt to move accordingly to its instincts.

The implemented MS is just a starting point to check our hypothesis on movement generation; actually there are some drawbacks, especially for the limited power of the State-Action table, and it misses some biologically aspects that has been proven to be important for human, such as the associative memory of sequence patterns.[25] But preliminary results confirm what we expected from the robot: the motor primitive paradigm is a good candidate for movement learning and generation, and the implemented cognitive architecture exploits well for the integration of perception and action for learning a sensorimotor map.

5. Conclusion

This paper deals with a novel approach to developmental robotics; it presents an intentional 'Goal- generating' architecture, namely IDRA, which simulates the structure and the interaction among three brain areas, the amygdala, the thalamus, and the cortex, which are known to play a key role in the human cognitive development, in terms of information processing, conceptual resources and perceptual skill. We integrate this architecture with a MS, which allows implementing the active sensing paradigm in combination with the motor primitives concept for the generation of new movements. In short, the three mentioned cerebral areas are translated with three corresponding modules, the WM for the cortex, the IM for the amygdala, and the GM for the thalamus. WM and GM are grouped in a fourth module, the DM, making the strict interaction between the thalamus and a single sub-area of the cortex. Several DMs are grouped in a layered open network, based on the principles of HOA: its input is coming from sensors and two outputs are sent to the MS, which is in charge of selecting and composing a movement to fulfill the actual goal, whether innate or acquired.

The proposed model wants to specifically consider basic sensing and actuation, two activities required for learning and for cognitive development. Our model is able to adapt to virtually any kind of sensory input, exploiting the statistic structure of data exactly as the human brain does; for actuation a simple motor primitive paradigm has been adopted.

The two sets of experiments here reported are designed to test the validity of the model, in particular its ability to learn new goals (as in the first experiment where the robot learns to distinguish shapes) and to compose new behaviors, consistent with these goals (as in the second experiment). These two skills compose the active sensing activity. The results show that the agent is able to take sensorial input and to learn how to compose movements for fulfilling its instincts.

Overall our results confirm the basic choices illustrated in the introduction; here we discuss some aspects of them with reference to other relevant literature.

The open-ended development of motor skills has attracted research in robotics due to the obvious role that the sensory-MS has in planning and action executing. As pointed out in [39] on the question about which internal motivations signals are best suited to decide which skills to learn, the best signals are those based on mechanisms that measure the improvement in the competence rather than the errors. Our conclusions are similar; even though we are not considering the classic action-planning paradigm, but only a preliminary phase where the action is taken to fulfill an innate instinct, errors are not directly used for learning where only the measure of the instinct fulfillment is considered.

Law et al. [40] presents a system based on internal motivation to acquire eye-arm reaching skills. We do not discuss here the reaching task, for which we have demonstrated in past research [41] that it can be obtained by a purely unsupervised learning method starting from a system that reconstructs images encoded with sparse-coding features. Even though a system that uses neural coding as basic representation could be integrated in IDRA we have opted for directly using the high-level representation of the neural functions. Learning in IDRA is only based on the capability of clustering incoming signals and using Hebbian rule. Other learning paradigms need to be integrated to consider other emerging properties of the biological system. We have introduced a neuronal-inspired RL in IDRA in a more advanced MS, where the Actor-critic architecture has been selected.[42]

An interesting extension of the cognitive architecture is about language acquisition. Intrinsic motivation has been the driver in the model of [43] of the initial development of speech in infants. We are also experimenting IDRA in learning simple words from the babbling phase, considering this task as a special case of motor primitives.

In the investigation about the possible role of IDRA in defining higher level functions, we have already approached the categorization problem and experimented how to cascade different DMs to recognize images and sounds.[21]

Several more concepts will be further explored in the future. An important aspect for long-life learning is neural plasticity, a property of the human brain related to the endless processing of sensorial input necessary for cognitive development. In IDRA, we can maintain an active network with a working memory of the agent state, its motivations, its skills, but we do not have a forgetting mechanism needed to reduce the complexity of the network.

Note

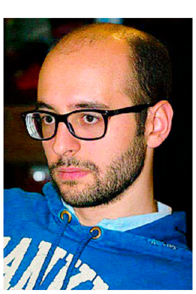
1. The log-polar images allow faster sampling rates on artificial vision systems without reducing the size of the field of view and the resolution on the central part of the retina.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Alessio Mauro Franchi graduated with a master's degree in computer science engineering from Politecnico di Milano, Italy, in 2013; he specialized in Artificial Intelligence and Robotics and his master thesis focuses on cognitive goal-generating architectures. He is currently a PhD student in Artificial Intelligence and Robotics at DEIB department, Politecnico di Milano, and collaborates with professor Gini; his activity concerns bioinspired artificial system with a particular interest for learning by demonstration and cognitive architectures.



Flavio Mutti received the BSc degree in computer engineering from Politecnico di Milano, Italy (2006), the MSc degree (summa cum laude) in Artificial intelligence and Robotics from the Politecnico di Milano, Italy (2009), and the PhD degree in Robotics and Artificial Intelligence, from the same institution, in 2013. His research focuses on bio inspired systems based on neuroscientific findings. His activity is focused on both low-level computational mechanisms, such as visuomotor mapping, and cognitive systems. He is currently co-founder of the B10NIX S.r.l., a company developing wearable interactive systems.

Giuseppina Gini graduated in Physics from the Milan University and specialized in Artificial Intelligence at Stanford University. Since 1987 she is an associate professor of robotics at Politecnico di Milano. Her research interests are in pattern recognition and modeling, with applications to robotics and intelligent systems. Author of more than 200 technical papers and books, and director of more than 20 international research programs.

ORCID

Giuseppina Gini http://orcid.org/0000-0002-3876-8360

References

[1] Bjorklund DF. Childrens thinking: cognitive development and individual differences. Belmont (CA): Wadsworth Publishing; 2011.

[2] Matarić MJ. Behavior-based robotics as a tool for synthesis of artificial behavior and analysis of natural behavior. Trends Cogn. Sci. 1998;2:82–86.

[3] Lungarella M, Metta G, Pfeifer R, et al. Developmental robotics: a survey. Connect. Sci. 2003;15:151–190.

[4] Manzotti R, Moderato P. Intentional change, intrinsic motivations, and goal generation. Behav. Brain Sci. 2014;37:31–432.

[5] Chella A, Manzotti R. Agi and machine consciousness. In: Wang P, Goertzel B, editors. Theoretical foundations of artificial general intelligence. Vol. 4. Paris: Atlantis Press; 2012. p. 263–282.

[6] Dileep G. How the brain might work: a hierarchical and temporal model for learning and recognition [dissertation]. Santa Clara (CA): Stanford University, 2008.

[7] Barto A. Intrinsic motivation and reinforcement learning. In: Baldassarre G, Mirolli M, editors. Intrinsically motivated learning in natural and artificial systems. Springer: Berlin Heidelberg; 2013. p. 17–47.

[8] Baldassarre G, Mirolli M. Intrinsically motivated learning in natural and artificial systems. Berlin: Springer-Verlag; 2013.

[9] Fiore VG, Sperati V, Mannella F, et al. Keep focussing: striatal dopamine multiple functions resolved in a single mechanism tested in a simulated humanoid robot. Front. Psychol. 2014;5, Article No: 124. doi:10.3389/fpsyg.2014.00124

[10] Baldassare G, Stafford T, Mirolli M, et al. Intrinsic motivations and open-ended development in animals, humans, and robots: an overview. Front. Psychol. 2014;5, Article No: 985. doi:10.3389/fpsyg.2014.00985

[11] Suzuki T, Yano S, Suzuki K. Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on Motivation oriented action selection for understanding dynamics of objects. Vol. 107; 2008 September; p. 846–851. doi:10.1109/IROS.2008.4651185

[12] Settles B. Active learning literature survey. Report No.: 1648. Madison (WI): University of Wisconsin-Madison, 2010.
[13] Merrick KE. Intrinsic motivation and introspection in reinforcement learning. IEEE Trans. Auton. Mental Dev. 2012;4:315–329.
[14] Power WT. A general feedback theory of human behavior: part ii. Perceptual Motor Skills. 1960:11;309–323.
[15] Bajcsy R. Active perception. In: Proceedings of the IEEE; August; 1988. p. 966–1005.
[16] Schmidhuber J. Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. Connection Sci. 2006;18:173–187.
[17] Frank M, Leitner J, Stollenga M, et al. Curiosity driven reinforcement learning for motion planning on humanoids. Front. Neurorob. 2014;7, Article No: 25. doi: 10.3389/fnbot.2013.00025
[18] Manzotti R, Tagliasco V. From "behaviour-based" robots to "motivations-based" robots. Rob. Auton. Syst. 2005;51:175–190.
[19] Hart CB, Giszter SF. Modular premotor drives and unit bursts as primitives for frog motor behaviors. J. Neurosci. 2004;24:5269–5282.
[20] Mussa-Ivaldi FA, Bizzi E. Motor learning through the combination of primitives. Philos. Trans. Roy. Soc. London Ser. B, Biol. Sci. 2000;355:1755–1769.
[21] Gini G, Franchi AM, Ferrini F, et al. Bio-inspired classification in the architecture of situated agents. Intell. Auton. Syst. 2014;13:577–589.
[22] Manzotti R, Mutti F, Gini G, et al. Cognitive integration through goal-generation in a robotic setup. Adv. Intell. Syst. Comput. 2013;196:1755–1769.
[23] Alexander GE, Crutcher M, DeLong MR. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, prefrontal and limbic functions. Prog. Brain Res. 1990;85:119–46.
[24] Sporns O. Networks of the brain. Cambridge (MA): The MIT Press; 2010.
[25] Hawkins J, Blakeslee S. On intelligence. New York (NY): St. Martin's Griffin; 2004.
[26] Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature. 1996;381:607–609.
[27] Sharma J, Angelucci A, Sur M. Induction of visual orientation modules in auditory cortex. Nature. 2000;404:841–847.
[28] Sur M, Angelucci A, Sharma J. Rewiring cortex: the role of patterned activity in development and plasticity of neocortical circuits. J. Neurobiol. 1999;41:33–43.
[29] Hyvärinen A, Hoyera PO. Independent component analysis: algorithms and applications. Neural Netw. 2000;13:411–430.
[30] Hyvärinen A. Independent component analysis: recent advances. Philos. Trans. Roy. Soc. London A: Math. Phys. Eng. Sci. 2012;371, Article No: 1984. doi: 10.1098/rsta.2011.0534
[31] Sherman SM, Guillery RW. Exploring the thalamus. San Diego (CA): Academic Press; 2000.
[32] Modhaa DS, Singhb R. Network architecture of the long-distance pathways in the macaque brain. In: Proceedings of the National Academy of Sciences of the United States of America; June; 2010. p. 13485–13490. Washington, DC.
[33] Adolphs R, Spezio M. Role of the amygdala in processing visual social stimuli. Prog. Brain Res. 2006;156: 363–738.
[34] Duncan S, Barrett LF. The role of the amygdala in visual awareness. Trends Cogn. Sci. 2007;11:190–192.
[35] Albus JS. A theory of cerebellar function. Math. Biosci. 1971;10:25–61.
[36] Doyona J, Bellec P, Amseld R, et al. Contributions of the basal ganglia and functionally related brain structures to motor learning. Behav. Brain Res. 2009;199:61–75.
[37] Flash T, Hochner B. Motor primitives in vertebrates and invertebrates. Curr. Opin. Neurobiol. 2005;15:660–666.
[38] Franchi AM, Mutti F, Gini G. A bioinspired intentional architecture for situated agents to explore human cognitive abilities. In: *Proceedings of the Conference-Workshop Bio-inspired Robotics – IARP*; May; 2014. p. 37–43. Rome.
[39] Santucci VG, Baldassarre G, Mirolli M. Which is the best intrinsic motivation signal for learning multiple skills? Front. Neurorob. 2013;7, Article No: 22. doi: 10.3389/fnbot.2013.00022
[40] Law J, Shaw P, Lee M, et al. From saccades to grasping: a model of coordinated reaching through simulated development on a humanoid robot. IEEE Trans. Auton. Mental Dev. 2014;6:93–109.
[41] Mutti F, Gini G. Visuomotor mapping based on hering's law for a redundant active stereo head and a 3 dof arm. In: Proceedings of 7th international ICST Conference on BIO-Inspired models of NETworks, Information, and Computing Systems (BIONETICS 2012). Vol. 134; December; 2014. p. 63–74. Lugano.
[42] Franchi AM, Attuario D, Gini G. The organization of cortex-ganglia-thalamus to generate movements from motor primitives: a model for developmental robotics. In: Workshop AIRO, AI*IA Conference; September. Ferrara, Italy; 2015. p. 1–7.
[43] Moulin-Frier C, Sao MN, Pierre-Yves O. Self-organization of early vocal development in infants and machines: the role of intrinsic motivation. Front. Psychol. 2013;4, Article No: 1006. doi:10.3389/fpsyg.2013.01006

Appendix 1. Implementation

The biological model and the architecture here presented are implemented in the IDRA software; the main idea behind the realization of this system was to create readable and modular source code, so that users may 'Plug&Play' their own robots in few simple steps, regardless its type, its structure or the available sensors. The entire project is written in Microsoft .NET C#, with the use of the IDE Visual Studio 2012 (VS 2012).

We started from the definition of all the classes and the library needed, collecting them in an UML class diagram; a simplified schema is reported in Figure A1.

The flow of information in the architecture is fairly simple: it starts with the 'Body Class' retrieving data from all the enabled sensors installed on the robot; these signals are passed to the 'Filter Class', which elaborates raw data to extract meaningful information (e.g. colors, lines or edges); then the 'Intentional Architecture' class receives this filtered information; in here it is defined the network of DMs and the IM, and when a new input is ready the system starts the computation, making the information flow from one layer of the network to the following. The two final outputs are sent back to the 'Body Class', where a set of available 'Behaviours'

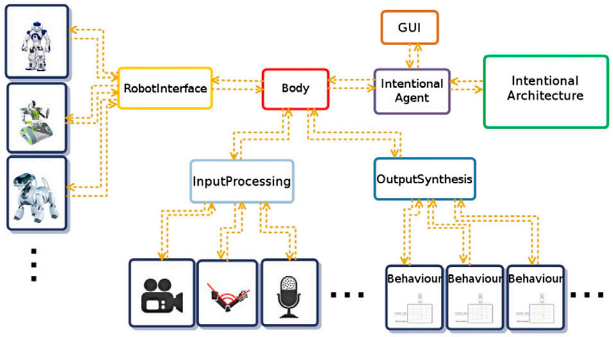


Figure A1. An abstraction of the UML diagram of the IDRA architecture.

(i.e. a list of actuators to be used) is defined and a proper movement is generated.

Going down into further details, the definitions of all the implemented robots are in the 'AgentsLib' folder; all the .cs files describing a new robot must respect the 'RobotInterface', which imposes the presence of all the necessary functions for this architecture to communicate with the agents (e.g. macros for connecting to robot or for reading sensors). All the data manipulation functions are defined in the 'InputLib' project; it contains an 'inputProcessing' class which instantiates all the filters necessary to deal with raw data; these filters are clustered into different classes, one for each type of sensor (e.g. audio, video, or tactile) and they contain specific filtering algorithm and methods for encapsulating data into predefined formats; this re-formatting of raw sensory data into an high-level structure is the key that makes the Intentional Architecture independent form the number and the type of sensors implemented on the robot. On the opposite side there is an 'OutputLib' project, which is necessary for the 'BodyLib' to know which actuators and behaviours are available on the robot in use. As soon as the Intentional Architecture has finished a cycle, the two outputs are sent to the 'outputSynthesis' class, which contains all the 'Behaviours' for the generation of robot's movements, each one with its list of actuators to be used and a State-Action table for the choice of the best movement. The 'outputSynthesis' class forwards its input to each Behavior, which returns the movements to apply to a specific actuator; a total list of movements is created and sent back to the Body class and then again to the Robot class for movement actuation.

The brain of the architecture is the 'IntentionalArchitectureLib', that is the actual implementation of the amydala-thalamo-cortical model previously presented; it contains the definition of all the modules: 'WorkingMemoryModule', 'DeliberativeModule', 'Goal-generatorModule' and 'InstinctsModule', which is in turn a collection of subclasses (the hard-coded instincts) each one dealing with a specific sensor type. The 'IntentionalArchitectureLib' is responsible for creating and managing the network of DMs, taking care of synchronization, data forwarding, inter-DMs communications and input/output retrieving.

All the software here detailed makes use of several XML files for saving and restoring all the configuration settings of the system; they are stored in a fixed directory structure and are editable both by hand and via the GUI for the most complex ones. In particular the 'Robot Folder' contains one XML file for each available robot, which describes for example its IP address for communication, its name, all the on board sensors and actuators.

Note that all the source code is available for free for testing purposes, collaboration or for experimenting with other robots. Please refer to the authors of this paper for more information.