CONAIM: A Conscious Attention-Based Integrated Model for Human-Like Robots

Alexandre da Silva Simões, *Member, IEEE*, Esther Luna Colombini, and Carlos Henrique Costa Ribeiro, *Member, IEEE*

***Abstract*—Understanding consciousness is one of the most fascinating challenges of our time. From ancient civilizations to modern philosophers, questions have been asked on how one is conscious of his/her own existence and about the world that surrounds him/her. Although there is no precise definition for consciousness, there is an agreement that it is strongly related to human cognitive processes such as attention, a process capable of promoting a selection of a few stimuli from a huge amount of information that reaches us constantly. In order to bring the consciousness discussion to a computational scenario, this paper presents conscious attention-based integrated model (CONAIM), a formal model for machine consciousness based on an attentional schema for human-like agent cognition that integrates: short- and long-term memories, reasoning, planning, emotion, decision-making, learning, motivation, and volition. Experimental results in a mobile robotics domain show that the agent can attentively use motivation, volition, and memories to set its goals and learn new concepts and procedures based on exogenous and endogenous stimuli. By performing computation over an attentional space, the model also allowed the agent to learn over a much reduced state space. Further implementation under this model could potentially allow the agent to express sentience, self-awareness, self-consciousness, autonoetic consciousness, mineness, and perspectivalness.**

***Index Terms*—Attention, cognition, machine consciousness, memory, robotics.**

I. Introduction

THROUGHOUT human history, it is possible to find reflections about the relation between body and mental phenomena. However, even after centuries of discussion, there is no common understanding about what consciousness is. For some authors, consciousness is everything we experience [1]. For others, it is the human ability to assign feelings to specialized mental capacities such as thinking, emotions, and speech [2]. Others define consciousness as the process that allows relevant information to remain online long enough so that it may be synchronously processed by multiple cortical networks [3]. Yet, some consider consciousness as the subjective experiences/awareness that one has when awake, following earlier suggestions that a more precise definition, given our currently inadequate scientific understanding of consciousness, is best left to the future [4]. Another understanding is that consciousness results from the complex interaction of the multimodal association areas of, at least, the parietotemporal and limbic cortices in an animal's brain, and refers to the capacity to experience oneself as a being subject to the past, present, and future, including the reflection on oneself as a being that is aware of its surrounding environment [5]. In general lines, there is a common understanding that consciousness is related to subjective experiences of humans, and that it is intrinsically connected to other human cognitive subsystems such as thinking, emotions, intention, decision-making, and reasoning.

Machine consciousness is the field of artificial intelligence concerned with the production of conscious processes in engineering hardware or software devices [6]. According to some definitions, a machine is considered conscious if besides the required mechanisms for perception, action, learning, and associative memory, it has a central executive that controls all the processes—conscious or subconscious—of the machine, and this central executive itself is driven by the machine's motivation and goal selection, attention-switching, semantic and episodic memory, using cognitive perception and cognitive understanding of motivations, thoughts, or plans to control learning, attention, motivations, and monitor actions [7]. A useful picture that places artificial consciousness among other concepts of the artificial intelligence domain—such as intelligence, life, and awareness—is shown in Fig. 1 [7], [8].

Several computational models have been proposed for machine consciousness (for an overview see [4], [9]). In this scenario, two main aspects stand out. First, one of the key concepts in consciousness models is attention. In this context, the recent proposition of an attentional model [10]–[12] capable of dealing with several types of attention—sustained, selective, oriented, divided, overt, and covert—and that can deal with endogenous and exogenous stimuli suggests further investigation on the applicability of this model for machine consciousness purposes. As a second relevant aspect, in the vast majority of consciousness models, authors present only a brief concept of the modules that compose the system, without relevant details about the connections and operation of these modules. This lack of formalization makes their implementation and extension impossible. In general lines, these works are also not followed by computational experiments that can demonstrate their capacities, preventing comparisons among them.

In order to bring the consciousness discussion to a computational scenario, this paper presents CONAIM, a novel conscious attention-based integrated model that comprises: short- and

Manuscript received February 01, 2015; revised September 08, 2015; accepted October 20, 2015. Date of publication January 14, 2016; date of current version September 27, 2017.

A. S. Simões is with the Department of Control and Automation Engineering, Campus of Sorocaba, Univ Estadual Paulista (UNESP), Sorocaba, Brazil (e-mail: assimoes@sorocaba.unesp.br).

E. L. Colombini is with the University of Campinas (UNICAMP), Campinas, Brazil (e-mail: esther.colombini@gmail.com).

C. H. C. Ribeiro is with the Technological Institute of Aeronautics (ITA), São José dos Campos, Brazil (e-mail: carlos@ita.br).

Digital Object Identifier 10.1109/JSYST.2015.2498542

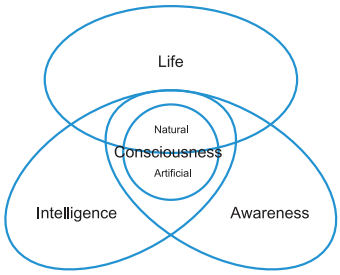

Fig. 1. Consciousness and its relation to intelligence, life, and awareness. Adapted from [7].

long-term memories, reasoning, planning, emotion, attention, decision-making, learning, motivation, and volition, and that is also able to accommodate other human-like cognitive processes. The formal model of the system is presented and, to the best of our knowledge, it represents the first known complete formalization of a model of cognitive processes that have been claimed to be linked to consciousness.

This paper is organized as follows. Section II introduces some of the basic concepts related to consciousness. Section III presents an overview of the computational models and theories for machine consciousness. Section IV discloses the proposed model for consciousness. Section V shows the experiments that were conducted. Section VI presents the experimental results and discussion. Finally, Section VII concludes this work.

II. Fundamentals of Consciousness

Some of the terms adopted in the consciousness literature have close and, sometimes, conflicting meanings. In this way, this section presents usual definitions of related concepts. A first important concept is being awake, i.e., not asleep. Humans can be awake without being conscious [13]. Another concept is sentience, the consciousness in the generic sense of simply being a sentient creature, one capable of sensing and responding to its world [14], [15], being capable of suffering or feeling pleasure or happiness [5]. Awareness usually refers to the process that occurs as a result of the interaction of the nervous system of the animal (including the sensory apparatus) and its environment, resulting in the ability of the animal to react to received stimuli. Awareness is, hence, related to terms such as sentience, perception, and cognition and does not require consciousness. A worm can be aware without being conscious [5]. Self-awareness is related to the being aware of yourself, i.e., the ability to introspect and to recognize itself as an individual separated from the environment and from other individuals.

Self-consciousness is a term of more complex definition. Some authors understand it as a synonym for self-awareness or consciousness in its minimalist sense [7]. However, it is possible to identify at least four distinct meanings for self-consciousness as follows [5]:

1) consciousness of its own body sensations and about its own movements;
2) ability to recognize itself in a mirror (e.g., ability demonstrated by chimpanzees, dolphins, and human beings);
3) consciousness of something happening in your environment and how this can affect you under a proper context;
4) consciousness of some decision and its consequences.

The autonoetic consciousness can be defined as the ability of a human being to, mentally, think himself/herself in the past and future, using his/her capacity of reasoning, anticipating situations, and reasoning about his/her own thoughts. Another important concept is volition, the cognitive process whereby an individual decides to perform a certain action, i.e., the mental activity through what intention is implemented [16].

III. Theories and Models for Machine Consciousness

One of the most accepted theories of consciousness is the integrated information theory (IIT) [1], [17], which claims that consciousness is integrated information, or, more precisely, that 1) the quantity of consciousness corresponds to the amount of integrated information (Φ) generated by a complex of elements and 2) the quality of experience is specified by the set of informational relationships generated within that complex. This theoretical framework can be applied to basic neurobiological observations as well as to artificial systems [18].

Computational models proposed for machine consciousness can be divided into four groups [4]: 1) models based on global workspace theory (GWT); 2) models based on internal self-models; 3) models based on higher level representations; and 4) models based on attention.

The GWT models are inspired by biological distributed and inter-communicating processes in regions of the cerebral cortex during mental efforts [19]–[21]. In these computational models, multiple specialized parallel processes cooperate and compete for the access to the global workspace. The coalition of winning processes can send information to all other processes through a broadcast system. The main hypothesis investigated by authors is that conscious experience can emerge through the collective interactions between the specialized processors via the global workspace. One of the first models for consciousness based on this paradigm was IDA [22], [23], a cognitive architecture that integrates several types of memory (perceptual, episodic, procedural, and working memory), consciousness, action selection, decision-making, and volition. One important contribution of IDA was the proposition of a cognitive cycle: perception, pre-conscious buffer, local association, competition for consciousness, broadcast, recruitment, setup of goal context hierarchy, action selection, and action execution. The adoption of a learning mechanism in IDA has led to the LIDA system [21]. Finally, the CERA-CRANIUM model [6], [24] presented a conscious system with attention and emotional learning based on the competition between specialized modules applied to a robotic agent.

Models based on internal self-models are biologically inspired by neural activation patterns that can be found in the human brain, and typically have an internal representation of the spatial properties of their bodies (a body image). The main hypothesis that guide authors in these works are [4]: 1) agents who reason about themselves can reason about their own reasoning about themselves [25]; 2) phenomenal consciousness could emerge from introspective reasoning mechanisms about perception [26]; and 3) some key properties of a self-concept

(existence, continuity over time, supervenience on a physical substrate, etc.) can form the basis for self-modeling in intelligent agents, regardless of whether or not they are embodied [27]. Some examples of the implementation of these models in robots can be found in the literature [28]–[30].

Models based on higher level representations consider that conscious mental states are distinguished from unconscious mental states by their level of representation. CLARION [31], [32] is an architecture with declarative and procedural knowledge, control mechanisms, and goals with two levels: 1) a higher conscious level—with a symbolic knowledge representation—and 2) a lower unconscious level—where concepts are represented by neural networks activity patterns. Another architecture with five levels was proposed in [33]. The lowest level is a reactive level, whereas the highest one—where conscious emerges in tasks that require cognitive efforts—is a symbolic rules level concerned with the robot movements. The CERA-CRANIUM—discussed in GWT models—can also be classified as a higher level model since its architecture has distinct cognitive levels. Similar works are [34] and [35].

At each moment, one is conscious of only a fraction of the information that reaches his/her sensory system [4]. Attention is the mechanism responsible for actively selecting, which stimulus will receive the attentional focus. Consciousness and attention are, therefore, extremely related processes [36]–[40]. Although there is such proximity, they are not, at least apparently, the same process [40], [41], since they have distinct functions and neural mechanisms [42]–[46] that occur jointly. Recent studies suggest that stimuli not attended by the attentional mechanism are not able to reach consciousness [3]. In this way, consciousness could be positioned at a distinct hierarchical level, and would be related to attention, competition, and learning [47]. Some authors argue that attention and consciousness are distinct mechanisms that feed the same decision process that leads to behavior [46].

Given the undeniable proximity between consciousness and attention, several computational models have been proposed based on attentional mechanisms. Even in these models, however, consciousness is interpreted in different ways. A computational model concerned with the visual perception composed of layers of spiking neurons was presented in [48]. These neurons were capable of building supermaps based on changing in their activities due to changing in input image features. According to authors, patterns in these maps could be used as a representation of consciousness. A distinct approach was proposed in [49]–[51] where several distinct interconnected modules without a precise hierarchical structure are capable of recruiting other modules in a competitive process to attend to the same topic they are working on. The machine, hence, could reach consciousness of a topic when several modules cooperate focusing their attention on the perceptions related to this topic. The model proposed by [7] and [52] presents an architecture composed of three main blocks: 1) a sensory-motor block; 2) episodic memory and learning block; and 3) executive central block. The attention shift results from the competition between several types of signals (internal thinking, new perceptions, shifts in motivation, etc.). A model that integrates consciousness, attention, and creativity is presented in [41]. Creativity could be used by subjects to amplify their subjective experiences, and is reached by decreasing the intensity of the attentional process, allowing the generation of unconscious thoughts until the subject can reach a mental state suitable for solving the problem.

IV. Proposed Model

This section presents: 1) the architecture of the proposed model and its main aspects; 2) model formalization; and 3) model operation.

A. Model Architecture

The proposed model consists of an architecture and a formal description of its structures and corresponding operations. The architecture is presented in Fig. 2 and is composed of two main systems.

1) An *attentional system*, following the *selection for perception* components of the model proposed in [11], comprises sensory memory, feature maps, weights associated with feature maps, combined feature maps, saliency map, and attentional map. The course of the attentional dynamics also follows the one proposed in [11].
2) A *cognitive system*, which replaces the *selection for action* component proposed by [11], composed of:
   a) *Decision-making* component that decides on the actions to be performed by the agent;
   b) *Short-term memory* component (working memory) that stores a small amount of information—typically received from the saliency map and sensory memory—which can be used by various modules of the cognitive agent system, in particular for long-term memory storage and recall;
   c) *Long-term memory* component stores, indefinitely, large amounts of information, consisting of:
      i) Episodic memory component that stores data, in a declarative form, relative to specific events, specifically *what*, *when*, and *where*, using a suitable data structure.
      ii) Semantic memory component that stores facts and knowledge about the world in a declarative way, using an appropriate data structure.
      iii) Procedural memory component that stores knowledge regarding procedures that can be executed by the agent, typically in the form of state-action pairs.
   d) *Evaluation* component: groups variables, lists, and functions related to the agent's performance, which composed of:
      i) Goals that stores the agent's current goal(s);
      ii) Evaluator: set of metrics designed to assess the agent's performance from different perspectives, typically acting over the internal state;
      iii) Evaluation state: the current values of the agent's state for all the metrics considered;
      iv) Task: current task under agent's execution;
      v) Volition that represents the process of transforming the agent's intention into a goal.

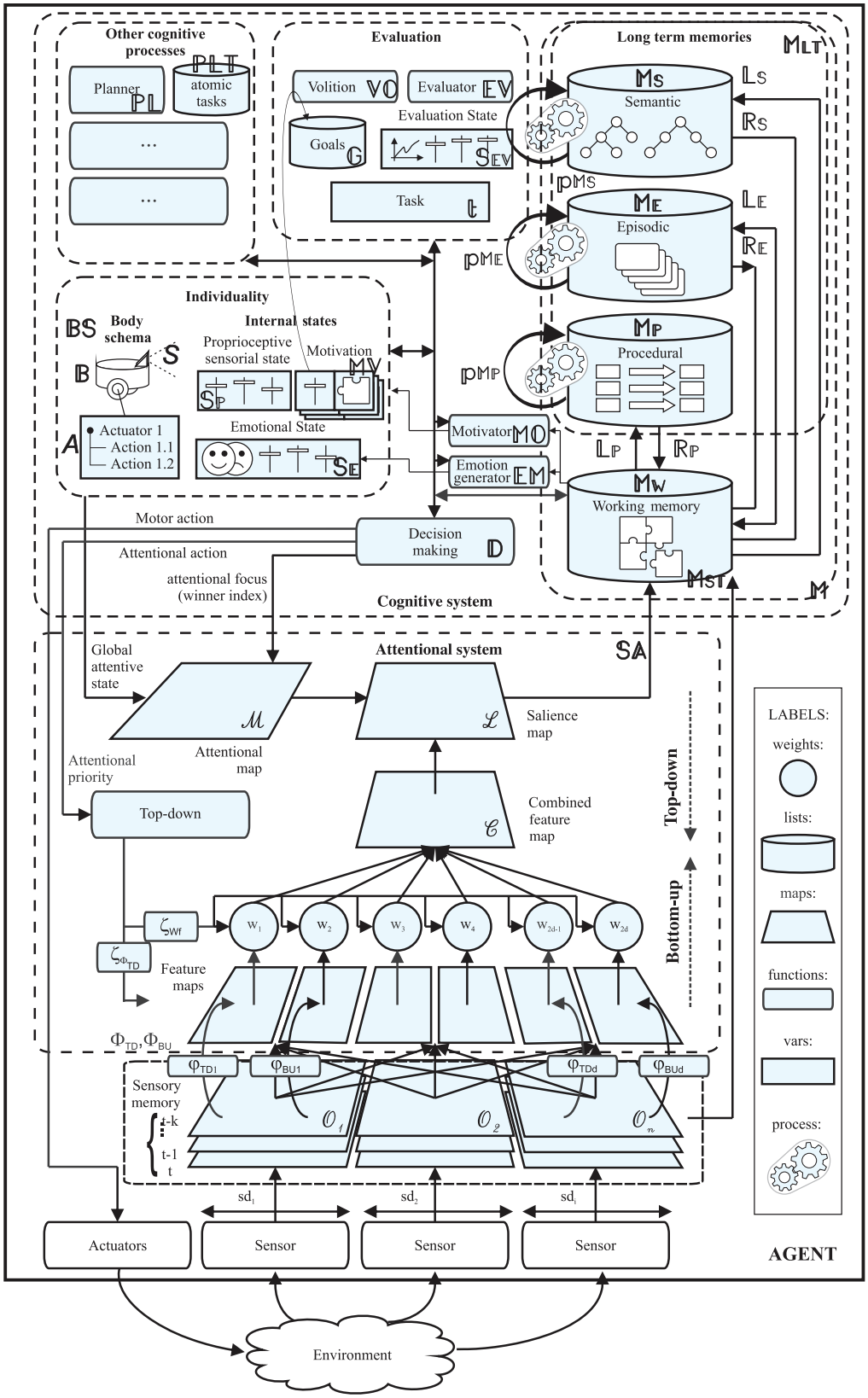

Fig. 2. CONAIM architecture, composed of an attentional system and a cognitive system: $\mathbb{M}$, memory; $\mathbb{M}_{\mathbb{LT}}$, long-term memory; $\mathbb{M}_{\mathbb{ST}}$, short-term memory; $\mathbb{M}_{\mathbb{S}}$, semantic memory; $\mathbb{M}_{\mathbb{E}}$, episodic memory; $\mathbb{M}_{\mathbb{P}}$, procedural memory; $\mathbb{M}_{\mathbb{W}}$, working memory; $\mathbb{D}$, decision-making; $\mathbb{EM}$, emotion generator; $\mathbb{MO}$, motivator; $\mathbb{BS}$, body schema; $\mathbb{B}$, body; $\mathbb{A}$, actuators, $\mathbb{S}$, sensors; $\mathbb{MV}$, motivation; $\mathbb{VO}$, volition; $\mathbb{EV}$, evaluator; $\mathbb{G}$, goals; $\mathbb{t}$, task; $\mathbb{PL}$, planner; $\mathbb{PLT}$: planner list; $\mathbb{p}$: processes; $\mathbb{S}_{\mathbb{E}}$: emotional state; $\mathbb{S}_{\mathbb{P}}$: proprioceptive state; $\mathbb{S}_{\mathbb{EV}}$: evaluation state; $\mathbb{S}_{\mathbb{A}}$: attentional state.

e) *Cognitive processes* that combine any number of the agent's cognitive functions, such as planning and state estimation;
f) *Individuality* component that composed of a set of inner states, such as proprioceptive, motivation, emotional state, and a body schema of the agent.

3) A set of *processes* with diverse purposes, running in background, which can act over the cognitive system components, not necessarily synchronized to its remaining components.

The system operates considering that any function or cognitive system process can request information from any other module and/or change the agent's internal states at any time. Therefore, the flow of information through the decision-making module is not mandatory.

B. Formal Model

This section presents a formal description of the proposed model. It uses the variables and functions defined in the *attentional system* formalized in [11]. The *cognitive system* can be described as follows.

1) A set of actions $\mathbb{A}$ with an arbitrary number of elements $\mathbb{a}_i$ termed *actions*, with $\{\mathbb{a}_1, \mathbb{a}_2, \ldots, \mathbb{a}_n\} \in \mathbb{A}$, and $\mathbb{A} = \mathbb{A}_\mathbb{A} \cup \mathbb{A}_\mathbb{M}$.
2) A set of attentional actions $\mathbb{A}_\mathbb{A}$ with an arbitrary number of elements $\mathbb{a}_{\mathbb{A}i}$ termed *attentional actions*, with $\mathbb{A}_\mathbb{A} = \{\mathbb{a}_{\mathbb{A}1}, \mathbb{a}_{\mathbb{A}2}, \ldots, \mathbb{a}_{\mathbb{A}n}\}$.
3) A set of motor actions $\mathbb{A}_\mathbb{M}$ with an arbitrary number of elements $\mathbb{a}_{\mathbb{M}i}$ termed *motor actions*, with $\mathbb{A}_\mathbb{M} = \{\mathbb{a}_{\mathbb{M}1}, \mathbb{a}_{\mathbb{M}2}, \ldots, \mathbb{a}_{\mathbb{M}n}\}$.
4) A list of *goals* $\mathbb{G}$ with an arbitrary number of elements $\mathbb{g}_i$, with $\mathbb{G} = \{\mathbb{g}_1, \mathbb{g}_2, \ldots, \mathbb{g}_n\}$.
5) A set of tasks $\mathbb{T}$ with an arbitrary number of elements $\mathbb{t}_i$ termed *tasks*, with $\mathbb{T} = \{\mathbb{t}_1, \mathbb{t}_2, \ldots, \mathbb{t}_n\}$, and the specific task $\mathbb{t}$ named agent's *current task*.
6) A set of states $\mathbb{S}$ with an arbitrary number of elements $\mathbb{s}_i$ termed *states*, where $\mathbb{S} = \{\mathbb{s}_1, \mathbb{s}_2, \ldots, \mathbb{s}_n\}$, with $\mathbb{S} = \mathbb{S}_\mathbb{A} \cup \mathbb{S}_\mathbb{E} \cup \mathbb{S}_{\mathbb{EV}} \cup \mathbb{S}_\mathbb{P}$.
7) A set of attentional states $\mathbb{S}_\mathbb{A}$ with an arbitrary number of elements such that $\mathbb{S}_\mathbb{A} = \{\mathbb{s}_{\mathbb{A}1}, \mathbb{s}_{\mathbb{A}2}, \ldots, \mathbb{s}_{\mathbb{A}n}\}$, where each element $\mathbb{s}_{\mathbb{A}i}$, denoted *attentional state*, consists of a *saliency map* of m dimension, termed $\mathcal{L} \in \Re^m$, with elements represented by $\{l_1, l_2, \ldots, l_m\}$, according to the model proposed by [11].
8) A set of emotional states $\mathbb{S}_\mathbb{E}$ with an arbitrary number of elements, where each term $\mathbb{s}_{\mathbb{E}i}$, denoted *emotional state*, is composed of an arbitrary number *emotional variables* $\mathbb{v}_\mathbb{E}$, i.e., $\mathbb{s}_\mathbb{E} = \{\mathbb{v}_{\mathbb{E}1}, \mathbb{v}_{\mathbb{E}2}, \ldots, \mathbb{v}_{\mathbb{E}n}\}$, where $\mathbb{v}_{\mathbb{E}i} \in \Re$.
9) A set of evaluation states $\mathbb{S}_{\mathbb{EV}}$ with an arbitrary number of elements, where each element $\mathbb{s}_{\mathbb{EV}i}$, denoted *evaluation state*, is constituted of an arbitrary number of *evaluation variables* $\mathbb{v}_{\mathbb{EV}}$, in other words, $\mathbb{s}_{\mathbb{EV}} = \{\mathbb{v}_{\mathbb{EV}1}, \mathbb{v}_{\mathbb{EV}2}, \ldots, \mathbb{v}_{\mathbb{EV}n}\}$, where $\mathbb{v}_{\mathbb{EV}i} \in \Re$.
10) A set of proprioceptive states $\mathbb{S}_\mathbb{P}$ composed of an arbitrary number of elements, with each element $\mathbb{s}_{\mathbb{P}i}$, termed *proprioceptive state*, constituted of an arbitrary number of *proprioceptive variables* $\mathbb{v}_\mathbb{P}$, i.e., $\mathbb{s}_\mathbb{P} = \{\mathbb{v}_{\mathbb{P}1}, \mathbb{v}_{\mathbb{P}2}, \ldots, \mathbb{v}_{\mathbb{P}n}\}$, with $\mathbb{v}_{\mathbb{P}i} \in \Re$.
11) A set of motivations $\mathbb{MV}$ of the agent with an arbitrary number of elements $\mathbb{mv}$ denoted *motivation*, where each element $\mathbb{mv}_i \in \mathbb{MV}$ is composed by a pair that contains $\mathbb{v}_{\mathbb{MV}} \in \Re$ denoted *motivation value*, and a memory element $\mathbb{m} \in \mathbb{M}$, in other words, $\mathbb{mv}_i = \{\mathbb{v}_{\mathbb{MV}i}, \mathbb{m}_i\}$.
12) A superset named *memory* $\mathbb{M}$, with $\mathbb{M} = \mathbb{M}_{\mathbb{LT}} \cup \mathbb{M}_{\mathbb{ST}}$.
13) A set named *short-term memory* $\mathbb{M}_{\mathbb{ST}}$, with $\mathbb{M}_{\mathbb{ST}} = \mathbb{M}_\mathbb{W}$.
14) A set named *long-term memory* $\mathbb{M}_{\mathbb{LT}}$, with $\mathbb{M}_{\mathbb{LT}} = \mathbb{M}_\mathbb{E} \cup \mathbb{M}_\mathbb{P} \cup \mathbb{M}_\mathbb{S}$.
15) A doubly linked list of elements named *working memory* $\mathbb{M}_W$ that can able to admit an arbitrary number of elements $\mathbb{m}_W$, i.e., $\mathbb{M}_\mathbb{W} = \{\mathbb{m}_{\mathbb{W}1}, \mathbb{m}_{\mathbb{W}2}, \ldots, \mathbb{m}_{\mathbb{W}n}\}$. The dynamics of insertions and removals in the list follows the first-in-first-out (FIFO) principle. It is recommended modeling elements $\mathbb{m}_W$ as proposed by [27].
16) A doubly linked list of elements named *semantic memory* $\mathbb{M}_\mathbb{S}$ that can comprise an arbitrary number of elements $\mathbb{m}_\mathbb{S}$, with $\mathbb{M}_\mathbb{S} = \{\mathbb{m}_{\mathbb{S}1}, \mathbb{m}_{\mathbb{S}2}, \ldots, \mathbb{m}_{\mathbb{S}n}\}$, where each element of the semantic memory $\mathbb{m}_{\mathbb{S}i}$ contains an arbitrary number of n-tuples d_{MS}, such that $\mathbb{m}_{\mathbb{S}i} = \{d_{\text{MS}1}, d_{\text{MS}2}, \ldots, d_{\text{MS}n}\}$. The ordering of the elements $\mathbb{m}_\mathbb{S}$ in the list $\mathbb{M}_\mathbb{S}$ is given by the value of $d_{\text{MS}i}$ denoted *semantic memory index element*.
17) A doubly linked list of elements named *episodic memory* $\mathbb{M}_\mathbb{E}$ that can able to admit an arbitrary number of elements $\mathbb{m}_\mathbb{E}$, with $\mathbb{M}_\mathbb{E} = \{\mathbb{m}_{\mathbb{E}1}, \mathbb{m}_{\mathbb{E}2}, \ldots, \mathbb{m}_{\mathbb{E}n}\}$, where each element of the episodic memory $\mathbb{m}_{\mathbb{E}i}$ contains an arbitrary number of n-tuples d_{ME}, such that $\{d_{\text{ME}1}, d_{\text{ME}2}, \ldots, d_{\text{ME}n}\} \in \mathbb{m}_{\mathbb{E}i}$. The ordering of the elements $\mathbb{m}_\mathbb{E}$ in the list $\mathbb{M}_\mathbb{E}$ is given by the increasing value of $d_{\text{ME}i}$ denoted *episodic memory index element*. It is recommended modeling the elements $\mathbb{m}_\mathbb{E}$ as proposed by [53].
18) A doubly linked list of elements termed *procedural memory* $\mathbb{M}_\mathbb{P}$ that can comprise an arbitrary number of elements $\mathbb{m}_\mathbb{P}$, with $\mathbb{M}_\mathbb{P} = \{\mathbb{m}_{\mathbb{P}1}, \mathbb{m}_{\mathbb{P}2}, \ldots, \mathbb{m}_{\mathbb{P}n}\}$, where each element of the procedural memory $\mathbb{m}_{\mathbb{P}i}$ is a triple composed of a *task* $\mathbb{t} \in \mathbb{T}$, a set of qualifiers $\mathbb{d}_{\mathbb{MP}}$ and a *state-action set* $\mathbb{C}_{\mathbb{SA}}$, with an arbitrary number of elements $\mathbb{c}_{\mathbb{SA}i} \in \mathbb{C}_{\mathbb{SA}}$, with each element $\mathbb{C}_{\mathbb{SA}}$ formed by state $\mathbb{s} \in \mathbb{S}$ and action $\mathbb{a} \in \mathbb{A}$ pairs, with $\mathbb{c}_{\mathbb{SA}i} = \{\mathbb{s}_i, \mathbb{a}_i\}$, such that $\mathbb{m}_{Pi} = \{\mathbb{t}_i, \mathbb{d}_{\mathbb{MP}}, \mathbb{C}_{\mathbb{S}i}\}$.
19) A *planning function* $\mathbb{PL}$ that maps a goal $\mathbb{g} \in \mathbb{G}$ to a set with an arbitrary number of tasks $\mathbb{t} \in \mathbb{PLT}$ with $\mathbb{PLT}$ represented by a list of tasks internal to the planner. In other words,

$$\mathbb{g} \xrightarrow{\mathbb{PL}} \{\mathbb{t}_1, \mathbb{t}_2, \ldots, \mathbb{t}_n\} \in \mathbb{PLT}.$$

20) A *volition function* $\mathbb{VO}$ that maps motivations $\mathbb{mv} \in \mathbb{MV}$ into goals $\mathbb{g} \in \mathbb{G}$ is given by

$$\{\mathbb{MV}\} \xrightarrow{\mathbb{VO}} \{\mathbb{G}\}.$$

21) An *evaluation function* $\mathbb{EV}$ that maps the agent's current state $\mathbb{S}$ its goals $\mathbb{G}$ and task $\mathbb{t} \in \mathbb{PLT}$ to the agent's evaluation state $\mathbb{S}_{\mathbb{EV}}$. In other words,

$$\{\mathbb{S}_{\mathbb{A}}, \mathbb{S}_{\mathbb{E}}, \mathbb{S}_{\mathbb{EV}}, \mathbb{S}_{\mathbb{P}}, \mathbb{G}, \mathbb{t}, \mathbb{PLT}\} \xrightarrow{\mathbb{EV}} \{\mathbb{S}_{\mathbb{EV}}\}.$$

22) A *motivation function* $\mathbb{MO}$ that maps the content of the working memory $\mathbb{M}_{\mathbb{W}}$ to the agent's motivation $\mathbb{MV}$, i.e., $\mathbb{MO}$: $\mathbb{M}_{\mathbb{W}} \rightarrow \mathbb{MV}$.

23) An *emotion generator* $\mathbb{EM}$ that maps the content of the working memory $\mathbb{M}_{\mathbb{W}}$ into an emotional state $\mathbb{S}_{\mathbb{E}}$ of the agent: $\mathbb{EM}$: $\mathbb{M}_{\mathbb{W}} \rightarrow \mathbb{S}_{\mathbb{E}}$.

24) A *decision-making function* $\mathbb{D}$ that typically coordinates all cognitive system lists and elements choosing an action $a \in A$, which can be a motor action $\mathbb{a}_{\mathbb{M}} \in \mathbb{A}_{\mathbb{M}}$ or an attentional action $\mathbb{a}_{\mathbb{A}} \in \mathbb{A}_{\mathbb{A}}$, and a number i denoted *attentional process winner index* or *attentional focus*. In other terms,

$$\{\mathbb{S}_{\mathbb{A}}, \mathbb{S}_{\mathbb{E}}, \mathbb{S}_{\mathbb{EV}}, \mathbb{S}_{\mathbb{P}}, \mathbb{G}, \mathbb{T}, \mathbb{M}_{\mathbb{W}}, \mathbb{MV}, \mathbb{PLT}, \mathbb{BS}\} \xrightarrow{\mathbb{D}}$$
$$\{\mathbb{a}_{\mathbb{M}}, \mathbb{a}_{\mathbb{A}}, i, \mathbb{G}, \mathbb{T}, \mathbb{M}_{\mathbb{W}}, \mathbb{MV}, \mathbb{S}_{\mathbb{E}}\}.$$

25) A *body schema* $\mathbb{BS}$ of the agent that is a n-tuple that comprises the agent's set of sensors $\mathcal{S}$ actuators $\mathcal{A}$, and a n-tuple $\mathbb{B}$ that contains a set of descriptors of the physical characteristics of the agent structure (weight, volume, center of gravity, external form, etc.), such that $\mathbb{BS} = \{\mathcal{S}, \mathcal{A}, \mathbb{B}\}$.

26) A set of unconscious asynchronous *processes* $\mathbb{P}$, where each $\{\mathbb{p}_1, \mathbb{p}_2, \ldots, \mathbb{p}_n\} \in \mathbb{P}$ is represented by a process responsible for performing a mapping from a specific system component to updates in the agent's state and/or lists

$$\{\mathbb{S}_{\mathbb{A}}, \mathbb{S}_{\mathbb{E}}, \mathbb{S}_{\mathbb{EV}}, \mathbb{S}_{\mathbb{P}}, \mathbb{MV}, \mathbb{M}, \mathbb{M}_{\mathbb{W}}, \mathbb{G}, \mathbb{T}\} \xrightarrow{\mathbb{P}}$$
$$\{\mathbb{S}_{\mathbb{A}}, \mathbb{S}_{\mathbb{E}}, \mathbb{S}_{\mathbb{EV}}, \mathbb{S}_{\mathbb{P}}, \mathbb{MV}, \mathbb{M}, \mathbb{M}_{\mathbb{W}}, \mathbb{G}, \mathbb{T}\}.$$

27) A *semantic learning* $\mathbb{L}_{\mathbb{S}}$ that can perform a mapping usually is given by

$$\{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{S}}\} \xrightarrow{\mathbb{L}_{\mathbb{S}}} \{\mathbb{m}_{\mathbb{S}}\}.$$

28) An *episodic learning* $\mathbb{L}_{\mathbb{E}}$ that can perform a mapping usually is given by

$$\{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{E}}\} \xrightarrow{\mathbb{L}_{\mathbb{E}}} \{\mathbb{m}_{\mathbb{E}}\}.$$

29) A *procedural learning* $\mathbb{L}_{\mathbb{P}}$ that can perform a mapping usually is given by

$$\{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{P}}\} \xrightarrow{\mathbb{L}_{\mathbb{P}}} \{\mathbb{m}_{\mathbb{P}}\}.$$

30) A *semantic recall* $\mathbb{R}_{\mathbb{S}}$ that can perform a mapping usually is given by

$$\{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{S}}\} \xrightarrow{\mathbb{R}_{\mathbb{S}}} \{\mathbb{m}_{\mathbb{W}}\}.$$

31) An *episode recall* $\mathbb{R}_{\mathbb{E}}$ that can perform a mapping usually is given by

$$\{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{E}}\} \xrightarrow{\mathbb{R}_{\mathbb{E}}} \{\mathbb{m}_{\mathbb{W}}\}.$$

32) A *procedure recall* $\mathbb{R}_{\mathbb{P}}$ that can perform a mapping usually is given by

$$\{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{P}}\} \xrightarrow{\mathbb{R}_{\mathbb{P}}} \{\mathbb{m}_{\mathbb{W}}\}.$$

C. Model Operation

The operation of the model proposed in Section IV-B is given by:

1) *Start a new attentional cycle:*
   a) Compute the selection for perception component of the model proposed by [11] until a *saliency map* $\mathcal{L}$ is available;
   b) Compute the *winner* of the competitive process, in other words, the index of the element with maximum saliency, according to

$$\text{winner} = \arg\left(\max\left(\mathcal{L}\right)\right).$$

2) *Start a new cognitive cycle:*
   a) Create a new element in the working memory $\mathbb{m}_{\mathbb{W}} \in \mathbb{M}_{\mathbb{W}}$, inserting in $\mathbb{m}_{\mathbb{W}}$ the index of winner, the observations $\mathcal{O}_{\text{winner}}$ and the saliency map $\mathcal{L}$, such that

$$\{\text{ winner}, \mathcal{O}_{\text{winner}}, \mathcal{L}\} \longrightarrow \mathbb{m}_{\mathbb{W}} \in \mathbb{M}_{\mathbb{W}}.$$

   b) Compute the *semantic learning* $\mathbb{L}_{\mathbb{S}}$ over $\mathbb{m}_{\mathbb{W}}$ aiming at the insertion or extension of the concepts in $\mathbb{M}_{\mathbb{S}}$, i.e., $\mathbb{L}_{\mathbb{S}} : \{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{S}}\} \rightarrow \mathbb{m}_{\mathbb{S}}$.
   c) Compute the *episodic learning* $\mathbb{L}_{\mathbb{E}}$ over $\mathbb{m}_{\mathbb{W}}$ aiming at the insertion or extension of the episodes in $\mathbb{M}_{\mathbb{E}}$, i.e., $\mathbb{L}_{\mathbb{E}} : \{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{E}}\} \rightarrow \mathbb{m}_{\mathbb{E}}$.
   d) Compute the *procedural learning* $\mathbb{L}_{\mathbb{P}}$ over $\mathbb{m}_{\mathbb{W}}$ aiming at the insertion or extension of the procedures in $\mathbb{M}_{\mathbb{P}}$, i.e., $\mathbb{L}_{\mathbb{P}} : \{\mathbb{m}_{\mathbb{W}}, \mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{P}}\} \rightarrow \mathbb{m}_{\mathbb{P}}$.
   e) Compute a *semantic recall* $\mathbb{R}_{\mathbb{S}}$ over $\mathbb{m}_{\mathbb{W}}$, inserting in $\mathbb{m}_{\mathbb{W}}$ the content recalled $\mathbb{m}_{\mathbb{S}}$ or *null*, i.e., $\mathbb{R}_{\mathbb{S}} : \{\mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{S}}, \emptyset\} \rightarrow \mathbb{m}_{\mathbb{W}}$. $\mathbb{R}_{\mathbb{S}}$ can be performed based on many distinct aspects $\mathbb{d}_{\mathbb{MS}}$ of elements $\mathbb{ms}$ (shape, color, smell, purpose, number of recalls, etc.). For each $\mathbb{d}_{\mathbb{MS}}$, a distinct matching algorithm must perform a full search in the memory elements.
   f) Compute an *episode recall* $\mathbb{R}_{\mathbb{E}}$ over $\mathbb{m}_{\mathbb{W}}$, inserting in $\mathbb{m}_{\mathbb{W}}$ the content recalled $\mathbb{m}_{\mathbb{E}}$ or *null*, i.e., $\mathbb{R}_{\mathbb{E}} : \{\mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{E}}, \emptyset\} \rightarrow \mathbb{m}_{\mathbb{W}}$. $\mathbb{R}_{\mathbb{E}}$ can be performed based on many distinct aspects $\mathbb{d}_{\mathbb{ME}}$ of elements $\mathbb{m}_{\mathbb{E}}$ (time of event, environment conditions, objects, actions performed, etc.). For each $\mathbb{d}_{\mathbb{ME}}$, a distinct matching algorithm must perform a full search in the memory elements.
   g) Compute a *procedure recall* $\mathbb{R}_{\mathbb{P}}$ over $\mathbb{m}_{\mathbb{W}}$, inserting in $\mathbb{m}_{\mathbb{W}}$ the content recalled $\mathbb{m}_{\mathbb{P}}$ or null, i.e., $\mathbb{R}_{\mathbb{P}} : \{\mathbb{M}_{\mathbb{W}}, \mathbb{M}_{\mathbb{P}}, \emptyset\} \rightarrow \mathbb{m}_{\mathbb{W}}$. $\mathbb{R}_{\mathbb{P}}$ can be performed based on

distinct qualifiers $\mathbb{d}_{\mathbb{MP}}$ of a task (purposes, restrictions, conditions, etc.). Some reasoning can be necessary to check the applicability of a procedure in a given situation to achieve an expected result.

h) Compute the *motivation function* $\mathbb{MO}$. Motivation functions can be proposed based on a large number of psychological theories [16]. Typically, the motivation is reduced when certain activity is carried out over a long time, or if it is perceived as boring, difficult, or if the agent has more enjoyable activities. It also depends on the agent as an individual, on its emotional state and memories or on its curiosity regarding new stimuli. If the item in $\mathbb{m}_{\mathbb{W}}$ indicates that:
   i) $\mathbb{M}_{\mathbb{S}}$, $\mathbb{M}_{\mathbb{E}}$, or $\mathbb{M}_{\mathbb{P}}$ have a concept, episode, or procedure under significant evolution that should be investigated by the agent, or
   ii) Elements recalled have relevant emotions associated, or
   iii) Elements are frequently recalled from memories

   then the agent is *motivated* to relate to the concept, episode or procedure by doing:
   - if $\mathbb{m}_{\mathbb{W}} \subset \mathbb{MV}$, then update the value of the associated motivation $\mathbb{v}_{\mathbb{MV}}$.
   - else, create a new item $\mathbb{mv}_i \in \mathbb{MV}$, insert a high value of motivation in $\mathbb{v}_{\mathbb{MV}i} \in \mathbb{mv}_i$ and insert $\mathbb{m}_{\mathbb{W}}$ in $\mathbb{m}_i \in \mathbb{mv}_i$.

   and also to *update* the values associated to motivations $\mathbb{m}_{\mathbb{W}} \subset \mathbb{MV}$ according to a time decay policy, eventually removing items from the list.

i) Compute the *Emotion Generator* $\mathbb{EM}$ based on $\mathbb{m}_{\mathbb{W}}$ (data observed and recalled memories) to update the agent's *emotional state* $\mathbb{S}_{\mathbb{E}}$ composed by a number of emotions $\mathbb{v}_{\mathbb{E}}$. Although a large range of basic emotions can be investigated, some candidates are anger, disgust, fear, happiness, and sadness, surprise [54], [55]. Emotion functions must affect the intensity of the emotions trying to mimic the complex biological regulation that takes place in the cortex, hippocampus, thalamus, and related regions of the human brain. For an overview, see [7], [56], [57].

j) Compute the agent's *volition function* $\mathbb{VO}$, evaluating $\mathbb{MV}$ and $\mathbb{G}$, to decide whether a new goal $\mathbb{g} \in \mathbb{G}$ should be defined for the agent. $\mathbb{VO}$ is assumed to be a process distinct from $\mathbb{MO}$, since $\mathbb{VO}$ is more related to a deliberative process. It can be divided into preactional and actional phases. The preactional phase is concerned with forming interpretations and planning when, where, and how to initiate the action. $\mathbb{VO}$ must choose considering the real possibility to accomplish the goal, the real benefits on achieving it, environmental conditions, strength of motivation, etc. The actional phase is related to the action's execution evaluation and the reasons to sustain it. Some factors must be considered: ability to resist distractions, environment capacity to capture agent's attention, unpredicted obstacles, discouragement from performing action, emotional state [16].

k) If needed, execute the *planning function* $\mathbb{PL}$ to determine the atomic tasks $\{\mathbb{t}_1, \mathbb{t}_2, \ldots, \mathbb{t}_n\} \in \mathbb{PLT}$ that are required to fulfil the agent's goal $\mathbb{G}$

$$\mathbb{G} \xrightarrow{\mathbb{PL}} \{\mathbb{t}_1, \mathbb{t}_2, \ldots, \mathbb{t}_n\} \in \mathbb{PLT}.$$

l) Evaluate the goals $\mathbb{G}$, the agent's current task $\mathbb{t}$, its state $\mathbb{S}$ and the tasks that should be executed $\mathbb{PLT}$, and compute the agent's *evaluation state* $\mathbb{S}_{\mathbb{D}}$ through

$$\{\mathbb{S}_{\mathbb{A}}, \mathbb{S}_{\mathbb{EV}}, \mathbb{S}_{\mathbb{E}}, \mathbb{S}_{\mathbb{P}}, \mathbb{G}, \mathbb{t}, \mathbb{PLT}\} \xrightarrow{\mathbb{EV}} \mathbb{S}_{\mathbb{EV}}.$$

m) Execute a *decision-making* $\mathbb{D}$ function by doing:
   i) Considering the list of goals $\mathbb{G}$, the current agent's task $\mathbb{t}$, the agent's general state $\mathbb{S}$, and the tasks that should be executed $\mathbb{PLT}$, decide on the maintenance or *modification of the current task* $\mathbb{t}$. In other words, compute

$$\{\mathbb{S}_{\mathbb{A}}, \mathbb{S}_{\mathbb{E}}, \mathbb{S}_{\mathbb{EV}}, \mathbb{S}_{\mathbb{P}}, \mathbb{G}, \mathbb{t}, \mathbb{PLT}\} \xrightarrow{\mathbb{D}} \mathbb{t}.$$

   ii) Insert $\mathbb{t}$ as a working memory item $\mathbb{m}_{\mathbb{W}}$ and wait for a recall of the $\mathbb{R}_{\mathbb{P}}$ task from $\mathbb{M}_{\mathbb{P}}$, by doing:
   - If the procedure to solve the task is known, i.e., if $\mathbb{R}_{\mathbb{P}}$: $\{\mathbb{m}_{\mathbb{P}}, \mathbb{MP}\} \rightarrow \mathbb{m}_{\mathbb{W}}$ results in a $\mathbb{m}_{\mathbb{W}} \neq \varnothing$ then execute the task $\mathbb{t}$ using the procedure described by the state-action pairs $\mathbb{C}_{\mathbb{SA}} \in \mathbb{m}_{\mathbb{W}}$.
   - Otherwise, if the recall returns a *fail*, i.e., $\mathbb{m}_{\mathbb{W}} = \varnothing$, then evaluate if the agent should start a learning procedure for the task $\mathbb{t}$. If yes, a new element $\mathbb{mp}_i \in \mathbb{M}_{\mathbb{P}}$ with $\mathbb{mp}_i = \{\mathbb{t}_i, \mathbb{C}_{\mathbb{SA}i}\}$ should be created and the learning should be started.

3) Start a new attentional cycle (step 1) for time = time + 1.

Besides the model operation described, at least three unconscious asynchronous processes can be running in background.

a) A process $\mathbb{p}_{\mathbb{M}_{\mathbb{S}}}$ that permanently scans the semantic memory $\mathbb{M}_{\mathbb{S}}$ seeking to derive or to unify concepts, with a mapping typically represented by: $\mathbb{p}_{\mathbb{M}_{\mathbb{S}}}$: $\mathbb{M}_{\mathbb{S}} \rightarrow \mathbb{M}_{\mathbb{S}}$.
b) A process $\mathbb{p}_{\mathbb{M}_{\mathbb{E}}}$ that permanently scans the episodic memory $\mathbb{M}_{\mathbb{E}}$ seeking to derive or to unify episodes, with a mapping typically represented by: $\mathbb{p}_{\mathbb{M}_{\mathbb{E}}}$: $\mathbb{M}_{\mathbb{E}} \rightarrow \mathbb{M}_{\mathbb{E}}$.
c) A process $\mathbb{p}_{\mathbb{M}_{\mathbb{P}}}$ that permanently scans the procedural memory $\mathbb{M}_{\mathbb{P}}$ seeking to derive or to unify processes, with a mapping typically represented by: $\mathbb{p}_{\mathbb{M}_{\mathbb{P}}}$: $\mathbb{M}_{\mathbb{P}} \rightarrow \mathbb{M}_{\mathbb{P}}$.

V. Materials and Methods

A. Simulation Environment

The system modules were implemented in C++ with wxWidgets and the experiments were conducted on the USARSim

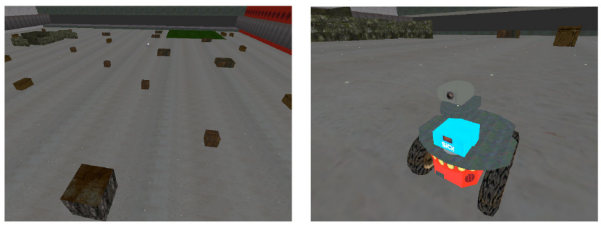

Fig. 3. Screenshots of the USARSim platform. (a) Overview of the DM-TestRoom250 scenario, with boxes of distinct sizes, walls, and victims, with wide open space. (b) Pioneer P2-AT virtual robot.

platform [58]. A P2AT simulated mobile robot with eight sonars and range scanner, both covering the frontal 180° of the robot was deployed in the environment. Sonars and laser have, respectively, 15 cm–5 m and 20 m range. All sensors include a random noise generator, configured to add up to 5% noise with a standard deviation $\sigma = 0.1$. The simulation scenario, shown in Fig. 3, was DM-Testroom-250, with large clear spaces, walls, irregular ground, and large fixed boxes. The robot was initially placed close to a box.

B. Experiment Description

Two distinct experiments were performed. EXP01 is an experiment where a robot with the attentional system was placed in the environment in front of a box. No cognitive system was adopted, and a simple controller drives the robot to the direction of the most salient stimulus. In EXP02, the robot was placed in the environment and controlled by the proposed conscious model, the robot was free to decide its own trajectory and to learn. The aim of this experiment is to show that consciousness-related aspects can emerge in this architecture due to the fact that a central executive $\mathbb{D}$ can coordinate learning, attention, motivation, and actions driven by its own motivations, goals, perception, attention-shifting, and memory-stored data. The attentional and cognitive systems setup as well as the system operation are detailed in the next sections.

C. Attentional System Setup

The observation spaces adopted were defined as follows. O_1 represents the observation space for sonar readings, defined by: $o_{1_{n_t}} = \text{sonar}_{n_t}$ with $n \in [1, 8]$; O_2 represents the observation space for range scanner readings, defined by: $o_{2_{n_t}} = \text{range}_{n_t}$ with $n \in [1, 180]$.

Three bottom-up feature maps were adopted.

1) $\mathcal{F}_1$ represents the level of discrepancy of the intensity of motion detected in the environment;
2) $\mathcal{F}_2$ represents the level of discrepancy of the direction of moving objects in the environment relative to the agent;
3) $\mathcal{F}_3$ represents the level of discrepancy of the distance between obstacles (static or moving) and the agent.

Details on the functions that perform the mapping from the observation spaces $\mathcal{O}$ to the feature maps spaces $\mathcal{F}$ can be found in [11]. Two top-down feature maps were adopted.

1) $\mathcal{F}_4$ represents the interest of the agent for elements disposed at a specific distance from it.
2) $\mathcal{F}_5$ represents the interest of the agent for specific regions of the space.

The saliency map $\mathcal{L}$ is an 8-dimensional state space and is used as the basis for the learning process instead of the observation space, composed by an 8-dimensional $\mathcal{O}_1$ and a 180-dimensional $\mathcal{O}_2$. This represents a 95% reduction in the usual state space.

D. Cognitive System Setup

In this experiment, a conscious robot was placed in the environment with the following configuration.

1) Actions set $\mathbb{A} = \mathbb{A}_{\mathbb{A}} \cup \mathbb{A}_{\mathbb{M}}$ with the following actions.
   a) $\mathbb{a}_{\mathbb{M}1}$ rotates toward the attentional focus.
   b) $\mathbb{a}_{\mathbb{M}2}$ moves forward.
   c) $\mathbb{a}_{\mathbb{A}1}$ drives the attentional focus through the top-down feature $\mathcal{F}_4$ to elements very close to the agent.
   d) $\mathbb{a}_{\mathbb{A}2}$ drives the attentional focus through the top-down feature $\mathcal{F}_4$ to elements close to the agent.
   e) $\mathbb{a}_{\mathbb{A}3}$ drives the attentional focus through the top-down feature $\mathcal{F}_4$ to elements far from the agent.
   f) $\mathbb{a}_{\mathbb{A}4}$ drives the attentional focus through the top-down feature $\mathcal{F}_5$ to the region on the agent extreme right.
   g) $\mathbb{a}_{\mathbb{A}5}$ drives the attentional focus through the top-down feature $\mathcal{F}_5$ to the region on the agent right side.
   h) $\mathbb{a}_{\mathbb{A}6}$ drives the attentional focus through the top-down feature $\mathcal{F}_5$ to the region on the agent front.
   i) $\mathbb{a}_{\mathbb{A}7}$ drives the attentional focus through the top-down feature $\mathcal{F}_5$ to the region on the agent left side.
   j) $\mathbb{a}_{\mathbb{A}8}$ drives the attentional focus through the top-down feature $\mathcal{F}_5$ to the region on the agent extreme left side.
2) Evaluation: List of goals $\mathbb{G}$ and Task $\mathbb{t}$, all initially empty.
3) Motivation $\mathbb{MV}$, initially empty.
4) Motivation function $\mathbb{MO}$: every new concept in $\mathbb{M}_{\mathbb{W}}$ that could not be located in the agent's semantic memory adds a motivation element in $\mathbb{MV}$ with a certain level of intensity that changes across time.
5) Semantic memory $\mathbb{M}_{\mathbb{S}}$: initially empty. Shape was adopted as the key element $\mathbb{d}_{\mathbb{MS}} \in \mathbb{ms}$. The object shape was initially represented using a point cloud. These points can typically evolve to a polygon using a *convex hull* algorithm performed by a semantic memory background process $\mathbb{pms}$, as soon as points close enough are available.
6) Semantic recall $\mathbb{R}_{\mathbb{S}}$: A full search in the semantic memory is performed looking for elements that have points close to those currently observed by the agent according to an Euclidian distance criteria.
7) Semantic learning $\mathbb{L}_{\mathbb{S}}$: the following rules were applied: i) if recall $\mathbb{RS}=\emptyset$ for $\mathbb{mw}$, then create a new entry $\mathbb{ms}$ in $\mathbb{MS}$ and store data $\mathbb{mw}$ in $\mathbb{d}_{MS} \in \mathbb{ms}$; ii) if object already exists in $\mathbb{ms}_i$, then append $\mathbb{mw}$ data in $\mathbb{ms}_i$.
8) Procedural memory $\mathbb{M}_{\mathbb{P}}$: initially empty.
9) Procedural recall $\mathbb{R}_{\mathbb{P}}$: a full search in the procedural memory is performed looking for a task $\mathbb{t}$ that can be applied over object $\mathbb{ms}$.

10) Procedural learning $\mathbb{L}_{\mathbb{P}}$: implemented through reinforcement learning with the following structure:
    a) algorithm: Q-learning;
    b) agent's state $\mathcal{L}$ and actions $\mathbb{A}$, as defined in Section V-D;
    c) state space discretization: five levels for each state (5^8 states);
    d) exploration rate: 0.95, decaying linearly with time to 0;
    e) learning rate: 0.9;
    f) temporal discount factor: 0.99;
    g) maximum number of trials: 100. Max 500 steps per trial;
    h) reward structure: $r = r + 10$ for each new data inserted into the concept under investigation in the semantic memory; $r = r + 3/d$, where d is the distance from the agent to the first element that generated the concept; $r = r - 2$ if the robot is close to collision; $r = r - 10$ if the robot already collided; $r = r - 50$ if the robot is permanently stuck.
11) Decision-making $\mathbb{D}$: The following rules were adopted: 1) if task $\mathbb{t}$ and goals list $\mathbb{G}$ are both empty, then drive the robot to the most salient stimulus; 2) if task $\mathbb{t}$ is empty and goals list $\mathbb{G}$ is not empty, then remove the highest priority goal $\mathbb{g}$ from $\mathbb{G}$, and make it the new task $\mathbb{t}$; 3) if task $\mathbb{t}$ requires a known procedure, then verify state $\mathbb{s}$ and execute corresponding action $\mathbb{a}$; 4) if task $\mathbb{t}$ requires an unknown procedure, then apply RL learning algorithm and execute referred action $\mathbb{a}$.

E. System Operation

The attentional and conscious courses are executed as depicted in Fig. 4. The resulting saliency map $\mathcal{L}$, the winner of the competitive process and its observations $\mathcal{O}_{\text{winner}}$ are stored in the working memory $\mathbb{M}_{\mathbb{W}}$. The arrival of new data in $\mathbb{M}_{\mathbb{W}}$ stimulates recalls and learning at procedural and semantic memories. Since memories were initially empty, no recalls are expected. In this case, semantic memory is expected to create new concepts based on the data received from the environment through the learning function $\mathbb{L}_{\mathbb{S}}$. One of the available mechanisms that can stimulate the agent to complete a concept is motivation. In this way, as soon as the first new data reaches $\mathbb{M}_{\mathbb{S}}$, a motivation is expected to be generated in $\mathbb{MV}$ to foster the investigation of this concept. The volition function $\mathbb{VO}$ typically transforms this motivation into a goal, adding it to the agent's goals list $\mathbb{G}$, which was initially empty. The decision-making function $\mathbb{D}$ is responsible for turning this goal into a task, by inserting it at the working memory $\mathbb{M}_{\mathbb{W}}$ to recall a procedure for this task. Since the procedural memory $\mathbb{M}_{\mathbb{P}}$ has no stored procedures, $\mathbb{D}$ can decide to start a new procedural learning $\mathbb{L}_{\mathbb{P}}$ for the task. This process is accomplished by choosing attentional actions $\mathbb{a}_{\mathbb{A}} \in \mathbb{A}_{\mathbb{A}}$ that can emphasize attention in the area close to the recent points through top-down actions. Learning can continue typically until the point cloud allow the formulation of the geometrical concept of the object. At each learning step, the state-action pairs are stored in the procedural memory as a recipe for the task of investigating a new object.

VI. Results and Discussion

The dynamics of the attentional system guided by CONAIM after the procedural learning is shown in Fig. 5. Fig. 5(m) shows the successive application of the top-down actions $\mathbb{a}_{\mathbb{a}5}$ and $\mathbb{a}_{\mathbb{a}7}$ (emphasize stimuli from the left and right sides of the robot) as decided by the cognitive system to guide the agent to investigate the box. The robot trajectory is shown in Fig. 5(a). Fig. 6 compares the stimulus-guided and the conscious-guided agent. Fig. 6(a) presents the trajectory of 100 trials from 8 distinct starting poses considering that the agent was always driven to the most salient stimulus. Fig. 6(b) presents the consciousness-guided agent trajectories during 100 learning trials from 8 distinct starting poses. The agent learned to walk around the object under interest collecting new information about it, until the geometrical concept of the box was formulated in the semantic memory. Fig. 6(c) and (d) presents the number of points collected from the box in both approaches.

A. Model Summary

As shown in the conducted experiment, the proposed model considers:

1) Working with *multiple sensors* $\mathcal{S}$ of different categories and dimensions, which are automatically merged by the sensory-attentional system [Fig. 5(n) and (o)].
2) *Attentiveness*, since the agent is able to react to environmental stimuli received through its sensors and modulated according to attentional and cognitive systems (Fig. 5).
3) *Volition* $\mathbb{VO}$ here modeled as a function able to turn motivations $\mathbb{mv} \in \mathbb{MV}$ into goals $\mathbb{g} \in \mathbb{G}$.
4) Bottom-up *exogenous* and top-down *endogenous* stimuli and the attentional course of selection;
5) The *sustained component of attention*, i.e., to maintain the focus of attention based on the agent's complete conscious cognitive scheme, now guided by the decision-making block ($\mathbb{D}$).
6) The *selective component of attention*, which raises to memory only a small subset of sensory information located in the vicinity of the attentional focus, dramatically reducing the sensory input space.
7) The *orienting component of attention*, which results on changing the focus of attention on the current cycle.
8) The *overt attention*, as the decision-making module $\mathbb{D}$ can directly adjust the sensory structures through the execution of a motor action $\mathbb{a}_{\mathbb{M}} \in \mathbb{A}_{\mathbb{M}}$.
9) The *covert attention*, as the decision-making module $\mathbb{D}$ can choose, if necessary, just for an attentional action $\mathbb{a}_{\mathbb{A}} \in \mathbb{A}_{\mathbb{A}}$, without a motor action $\mathbb{a}_{\mathbb{M}} \in \mathbb{A}_{\mathbb{M}}$ associated to it.

The proposed formal model points out the potential for further developments that could support:

1) *Sentience*, since the agent can have its emotional state variables $\mathbb{v}_{\mathbb{E}} \in \mathbb{S}_{\mathbb{E}}$ altered by the conscious experience;
2) *Self-awareness*, since the model can accommodate an individuality schema that contains internal states and body image, as well as planning, decision-making, and learning mechanisms that support insight and possible comparisons with other individuals;

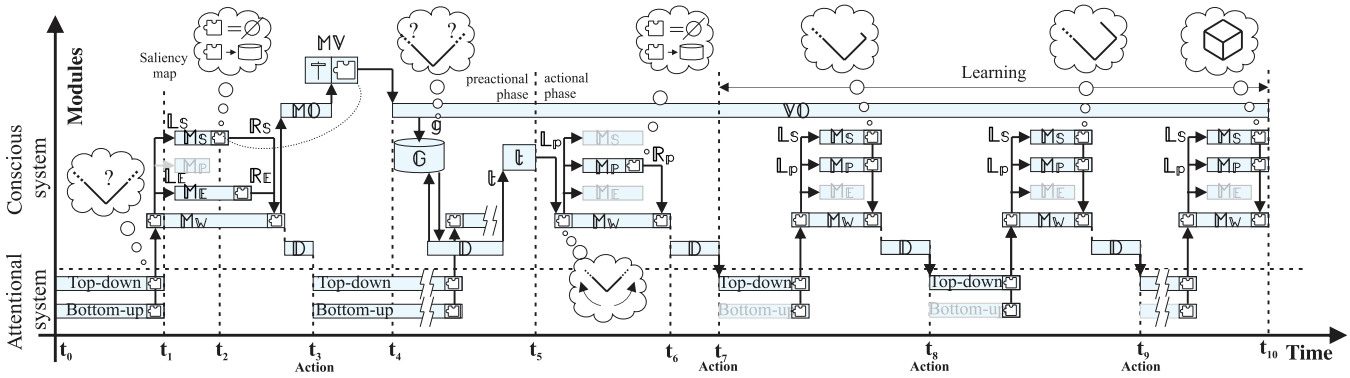

Fig. 4. CONAIM system dynamics during the experiment. $\mathbf{t}_0$: new stimuli arrive from the environment; $\mathbf{t}_1$ saliency map $\mathcal{L}$ is computed and working memory $\mathbb{M}_\mathbb{W}$ receives data regarding the new object; $\mathbf{t}_2$: recalls in memories are stimulated by $\mathfrak{mv}$. Semantic memory recall ($\mathbb{R}_\mathbb{S}$) fails indicating that this is a concept new to the agent. A new entry $\mathfrak{m}_\mathfrak{s}$ is created in $\mathbb{M}_\mathbb{S}$ with the collected data. $\mathbf{t}_3$: after memory recall ($\mathbb{R}_\mathbb{S}$ and $\mathbb{R}_\mathbb{E}$) fails, decision-making $\mathbb{D}$ drives the agent to the most salient stimulus. A new attentional cycle is started. Motivation function $\mathbb{MO}$ makes the agent interested on the new concept. $\mathbf{t}_4$: volition $\mathbb{VO}$ in preactional phase inserts a new goal $\mathfrak{g}$ at goals list $\mathbb{G}$, and decision-making $\mathbb{D}$ turns it into the new agent task $\mathfrak{t}$; $\mathbf{t}_5$: in order to investigate the new concept, the agent looks for an investigation procedure in procedural memory $\mathbb{M}_\mathbb{P}$, and the recall $\mathbb{R}_\mathbb{P}$ fails. A new entry $\mathfrak{m}_\mathfrak{p}$ is created in $\mathbb{M}_\mathbb{P}$; $\mathbf{t}_6$: decision-making $\mathbb{D}$ starts a new learning for $\mathfrak{m}_\mathfrak{p}$ by choosing attentional actions $\mathfrak{a}_\mathbb{A}$ that can emphasize stimuli in areas close to recently collected points (top-down influence); $\mathbf{t}_{7-10}$: for several interactions, this investigation procedure will be learned by the agent. New collected data about the concept will be stored in $\mathfrak{ms}$. The learning procedure of $\mathfrak{mp} \in \mathbb{M}_\mathbb{P}$ can continue until a complete concept $\mathfrak{ms} \in \mathbb{M}_\mathbb{S}$ can be considered fully formulated.

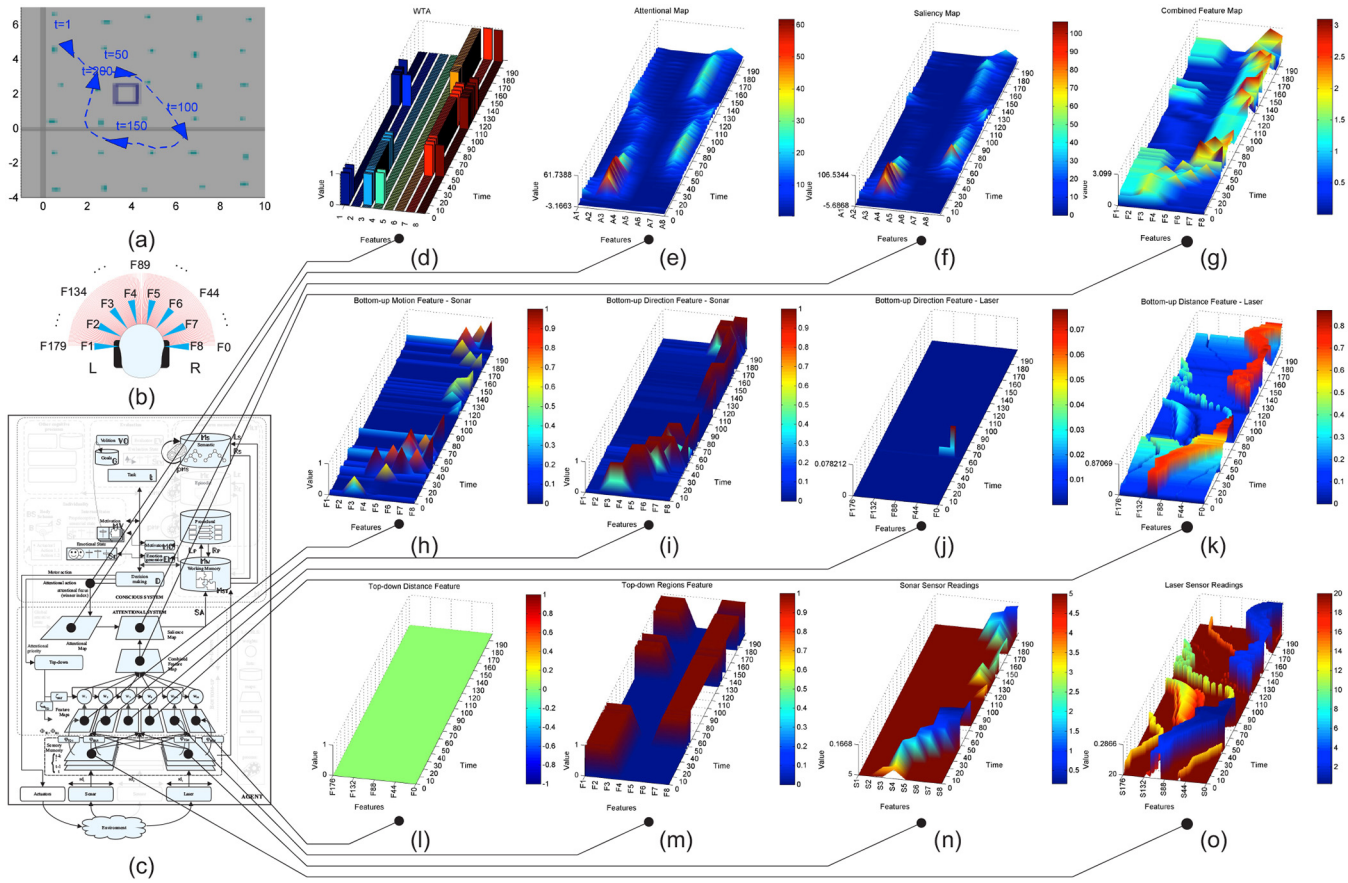

Fig. 5. Dynamics of the attentional system guided by the cognitive system in one of the experiments. (a) Trajectory of the agent around a box from $t = 0$ to $t = 200$ steps (dimensions in meters). (b) Disposition and labels for the eight sonar measurements [F1–F8] (blue) and 180 range scanner measurements [F0–F179] (red). (c) CONAIM architecture adopted in the the experiment (for details see Fig. 2). (d) Index of the *winner* of the attentional focus. (e) Attentional map $\mathcal{M}$. (f) Saliency Map $\mathcal{L}$. (g) Combined feature map $\mathcal{C}$. (h) *Bottom-up* feature map $\mathcal{F}_1$ (motion discrepancy). (i) *Bottom-up* feature map $\mathcal{F}_2$ sonar component (moving directions discrepancy). (j) *Bottom-up* feature map $\mathcal{F}_2$ laser component (moving directions discrepancy). (k) *Bottom-up* feature map $\mathcal{F}_3$ (distance discrepancy). (l) *Top-down* feature map $\mathcal{F}_4$ (interest on measurements with a specific distance). (m) *Top-down* feature map $\mathcal{F}_5$ (interest on specific regions). (n) Sonar readings $\mathcal{O}_1$ (meters). (o) Range scanner readings $\mathcal{O}_2$ (m). As shown in (m), after learning, the attentional actions $\mathfrak{a}_{\mathfrak{a}5}$ (drive attentional focus to right side) and $\mathfrak{a}_{\mathfrak{a}7}$ (drive the attentional focus to left side) are successively recovered from procedural memory $\mathbb{M}_\mathbb{P}$, guiding the robot to investigate the box.

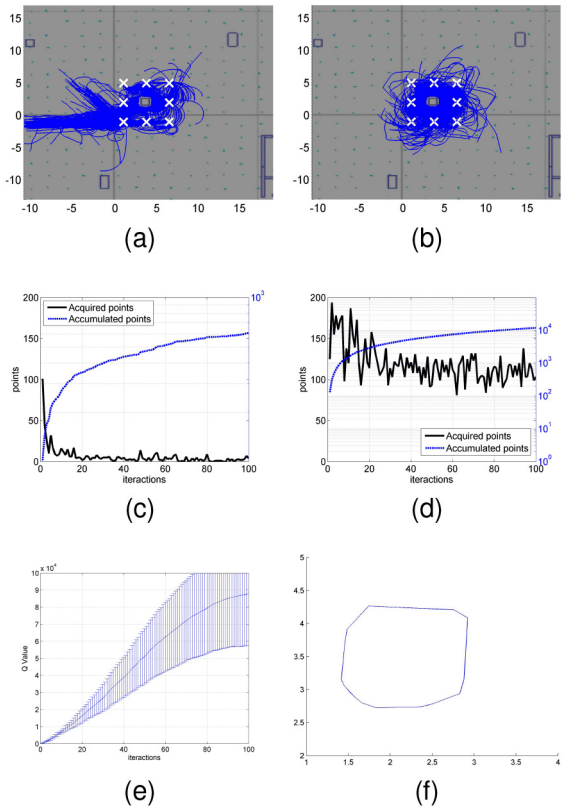

Fig. 6. Agent trajectories when exploring an environment with several boxes starting from eight distinct poses around a central box. (a) Trajectories of the robot guided by the most salient stimulus. (b) Trajectories of the robot guided by the CONAIM system performing the semantic and procedural learning. (c) and (d) Number of points of the central box acquired by the agents in "a" and "b," respectively, during computational steps and total amount of accumulated points. (e) Average Q values and standard deviation of the agent in "b" along learning trials for all experiments. (f) Example of resulting *convex hull* applied in the semantic memory of the agent in "b" over the point cloud.

3) *Autonoetic consciousness*, since the model can accommodate elements that enable reasoning about the past, present and future, especially when considering the modeling of the working memory proposed by [27];
4) *Mineness*, since objects frequently used by the agent can make use of one of its indicators $\mathbb{dms} \in \mathbb{Ms}$, in order to represent an ownership relationship;
5) *Perspectivalness*, since the mental mechanism can be expanded to reason about the perspective of the *self* compared to the other. The adoption of a point of view will depend on the capacity of building concepts in the semantic memory, and the agent's reasoning over them;
6) *Self-consciousness*, since the model can be expanded to allow the formulation of concepts about the surroundings of the agent, the assimilation of experiences and the analysis of the consequences of the agent decisions, usually based on learning and experience. The model could also allow the agent to be aware of its body, to recognize a similar and/or itself and to reason about the consequences of its decisions;
7) Other specialized modules such as language processing can be inserted in the other cognitive processes block.

VII. Conclusion

This paper proposed a novel model based on an attentional paradigm for machine consciousness consistent with various definitions of the concept available in the literature. The attentional system allows the agent to learn over the saliency map space instead of the traditional sensorial space, promoting a reduction of 95% on the original state-space. The conscious architecture includes cognitive elements such as memories, reasoning, planning, emotion, decision-making, learning, motivation, and volition. The model also provides a methodology for performing computation over these elements. Unlike other studies in the literature, where in most cases the authors briefly present the main concepts behind the modules that compose the system, this work details such modules, their connections, and their ways of operation. The formal model of the system is presented and, to our knowledge, it represents the first known complete formalization of a consciousness integrated model. The organization of the elements and processes involved in consciousness in an architecture also allows further comparison among distinct algorithms, as well as offers a base for testing consciousness, psychological, neuroscience, and machine learning theories, helping us to better understand human and machine consciousness. Experimental results suggest that this model is capable of supporting attentiveness, sentience, and self-awareness based on exogenous and endogenous stimuli. Although in these initial experiments only some basic components were engaged, there is room for the proposition of new functions (motivation, emotion, etc.) and modules that can far improve the agent behavior. Finally, we hypothesise that the model could be extended to demonstrate autonoetic consciousness, mineness, perspectivalness, and self-consciousness.

References

[1] G. Tononi, "An information integration theory of consciousness," *BMC Neurosci.*, vol. 5, no. 42, pp. 1–22, 2004.
[2] M. S. Gazzaniga, R. B. Ivry, and G. R. Mangun, Eds., *Cognitive Neuroscience: The Biology of Mind*. New York, NY, USA: Norton, 2002.
[3] M. A. Cohen, P. Cavanagh, M. M. Chun, and K. Nakayama, "The attentional requirements of consciousness," *Trends Cognit. Sci.*, vol. 16, pp. 411–417, Aug. 2012.
[4] J. A. Reggia, "The rise of machine consciousness: Studying consciousness with computational models," *Neural Netw.*, vol. 44, pp. 112–131, 2013.
[5] R. Arp, "Consciousness and awareness: Switched-on Rheostats: Response to de Quincey," *J. Conscious. Stud.*, vol. 14, no. 3, pp. 101–106, 2007.
[6] R. A. Moreno, A. L. Espino, and A. S. de Miguel, "Modeling consciousness for autonomous robot exploration," in *Proc. Int. Work-Conf. Interplay Nat. Artif. Comput. (IWINAC)*, 2007, vol. 4527, pp. 51–60.
[7] J. A. Starzyk and D. K. Prasad, "A computational model of machine consciousness," *Int. J. Mach. Conscious.*, vol. 3, pp. 255–281, 2011.
[8] M. Nisargadatta, *I Am That*. Bombay, India: Chetana, 1973.
[9] D. Gamez, "Progress in machine consciousness," *Conscious. Cognit.*, vol. 17, pp. 887–910, Sep. 2008.
[10] E. L. Colombini and C. C. Ribeiro, "An attentive multi-sensor based system for mobile robotics," in *Proc. IEEE Int. Conf. Syst. Man Cybern. (SMC)*, 2012, pp. 1509–1514.
[11] E. L. Colombini, *An attentional model for intelligent robotics agent*. Ph.D. dissertation, Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos, SP, Brazil, 2014.
[12] E. L. Colombini, A. S. Simões, and C. C. Ribeiro, "Top-down and bottom-up features combination for multi-sensor attentive robots," in *Proc. Int. Symp. Atten. Cognit. Syst. (ISACS); 23rd Int. Joint Conf. Artif. Intell. (IJCAI)*, Beijing, China, pp. 1–14, 2013.

[13] S. Laureys *et al.*, "Brain function in the vegetative state," *Acta Neurol. Belgica*, vol. 102, pp. 177–185, Dec. 2002.
[14] R. Van Gulick, "Consciousness," in *The Stanford Encyclopedia of Philosophy*, Zalta E. N., Ed. Stanford, CA, USA: Stanford Univ., Spring 2014.
[15] D. Armstrong, *What is Consciousness? In the Nature of Mind*. Ithaca, NY, USA: Cornell Univ. Press, 1981.
[16] J.-P. Broonen *et al.*, "Is volition the missing link in the management of low back pain?," *Joint Bone Spine*, vol. 78, pp. 364–367, Jul. 2011.
[17] G. Tononi, "Consciousness as integrated information: A provisional manifesto," *Biol. Bull.*, vol. 215, pp. 216–242, 2008.
[18] D. Gamez, "Information integration based predictions about the conscious states of a spiking neural network," *Conscious. Cognit.*, vol. 19, pp. 294–310, Mar. 2010.
[19] B. J. Baars, *A Cognitive Theory of Consciousness*. Cambridge, U.K.: Cambridge Univ. Press, 1988.
[20] B. Baars, "The conscious aspects hypothesis," *Trends Cognit. Sci.*, vol. 6, pp. 47–52, 2002.
[21] B. Baars and S. Franklin, "Conscious is computational: The LIDA model," *Int. J. Mach. Conscious.*, vol. 1, pp. 23–32, 2009.
[22] S. Franklin, A. Kelemen, and L. McCauley, "IDA: A cognitive agent architecture," in *Proc. IEEE Int. Conf. Syst. Man Cybern. (SMC)*, 1998, pp. 2646–2651.
[23] S. Franklin *et al.*, "Lida: A computational model of global workspace theory and developmental learning," in *Proc. AAAI Fall Symp. AI Conscious. Theor. Found. Curr. Approaches*, 2007, pp. 1–6.
[24] R. Arrabales, A. Ledezma, and A. Sanchis, "Towards conscious-like behaviour in computer game characters," in *Proc. IEEE Symp. Comput. Intell. Games*, 2009, pp. 217–224.
[25] D. Perlis, "Consciousness as self-function," *J. Conscious. Stud.*, vol. 4, pp. 509–525, 1997.
[26] D. McDermott, "Artificial intelligence and consciousness," in *Cambridge Handbook of Consciousness*, Cambridge, U.K.: Cambridge Univ. Press, 2007, pp. 117–150.
[27] A. Samsonovich and L. Nadel, "Fundamental principles of mechanisms of the conscious self," *Cortex*, vol. 41, pp. 669–689, 2005.
[28] O. Holland, "A strong embodied approach to machine consciousness," *J. Conscious. Stud.*, vol. 14, pp. 97–110, 2007.
[29] J. Takeno, "A robot succeds in 100% mirror image cognition," *Int. J. Smart Sens. Intell. Syst.*, vol. 1, pp. 891–911, 2008.
[30] J. Takeno, *Creation of a Conscious Robot*. Singapore: Pan Stanford, 2013.
[31] R. Sun, "Accounting for the computational basis of consciousness," *Conscious. Cognit.*, vol. 8, pp. 529–565, 1999.
[32] R. Sun and S. Franklin, "Computational models of consciousness," in *Cambridge Handbook of Consciousness*, P. Zelazo and M. Moscovitch, Eds., Cambridge, U.K.: Cambridge Univ. Press, 2007, pp. 151–174.
[33] T. Kitamura, T. Tahara, and K. Asami, "How can a robot have consciousness?," *Adv. Robot.*, vol. 14, pp. 263–275, 2000.
[34] A. Chella, "Towards robot conscious perception," in *Artificial Consciousness*, A. Chella and R. Manzotti, Eds., New York, NY, USA: Academic, 2007, pp. 124–140.
[35] A. Chella, M. Frixione, and S. Gaglio, "A cognitive architecture for robot self-consciousness," *Artif. Intell. Med.*, vol. 44, pp. 147–154, 2008.
[36] M. I. Posner and S. Dehaene, "Attentional networks," *Trends Neurosci.*, vol. 17, pp. 75–79, 1994.
[37] M. Velmans, *The Science of Consciousness*. Evanston, IL, USA: Routledge, 1996.
[38] P. M. Merikle and S. Joordens, "Parallels between perception without attention and perception without awareness," *Conscious. Cogn.*, vol. 6, pp. 219–236, 1997.
[39] J. K. O'Regan and A. Noe, "A sensorimotor account of vision and visual consciousness," *Behav. Brain Sci.*, no. 24, pp. 939–973, 2001.
[40] C. Koch and N. Tsuchiya, "Attention and consciousness: Two distinct brain processes," *Trends Cognit. Sci.*, vol. 11, pp. 16–22, Jan. 2007.
[41] J. G. Taylor, "The creativity effect: Consciousness versus attention," in *Proc. Int. Joint Conf. Neur. Netw. (IJCNN)*, 2010, pp. 1–8.
[42] C. Koch, *The Quest for Consciousness: A Neurobiological Approach*. Greenwood Village, CO, USA: Roberts and Publishers, 2004.
[43] V. G. Hardcastle, "Attention versus consciousness: A distinction with a difference," *Cognit. Stud. Bull. Jpn. Cognit. Sci. Soc.*, vol. 4, pp. 56–66, 1997.
[44] V. A. Lamme, "Why visual attention and awareness are different," *Trends Cognit. Sci.*, vol. 7, pp. 12–18, 2003.
[45] G. F. Woodman and S. J. Luck, "Dissociations among attention, perception, and awareness during object-substitution masking," *Psychol. Sci.*, vol. 14, pp. 605–611, 2003.
[46] C. Tallon-Baudry, "On the neural mechanisms subserving consciousness and attention," *Front. Psychol.*, vol. 2, pp. 397–397, Jan. 2012.
[47] S. Grossberg, "Linking attention to learning, expectation, competition, and consciousness," Dep. of Cog. and Neur. Syst., Boston Univ Center for Adap. Sys., Tech. Rep. CAS/CNS-2003-007, Boston, Apr. 2010.
[48] C. J. Tinsley, "Using topographic networks to build a representation of consciousness," *Biosystems*, vol. 92, pp. 29–41, Jan. 2008.
[49] P. Haikonen, *The cognitive approach to conscious machines*. New York, NY, USA: Academic, 2003.
[50] P. O. A. Haikonen, *Robot Brains. Circuits and Systems for Conscious Machines.* Hoboken, NJ, USA: Wiley, 2007.
[51] P. Haikonen, "Essential issues of conscious machines," *J. Conscious. Stud.*, vol. 14, pp. 72–84, 2007.
[52] J. A. Starzyk and D. K. Prasad, "Machine consciousness: A computational model," *Brain-Inspired Cognit. Syst. (BICS)*, Madrid, Spain, Jul. 14–16, 2010, pp. 428–439.
[53] D. Stachowicz and G. M. Kruijff, "Episodic-like memory for cognitive robots," *IEEE Trans. Auton. Ment. Develop.*, vol. 4, no. 1, pp. 1–16, Mar. 2012.
[54] P. Ekman, "Basic emotions," in *Handbook of Cognition and Emotion*, Dalgleish T. and Power M., Eds. Hoboken, NJ, USA: Wiley, 1999.
[55] J. Panksepp, "Affective consciousness: Core emotional feelings in animals and humans," *Conscious. Cognit.*, vol. 14, pp. 30–80, 2005.
[56] T. Ziemke and R. Lowe, "On the role of emotion in embodied cognitive architectures: From organisms to robots," *Cognit. Comput.*, vol. 1, pp. 104–117, 2009.
[57] A. R. Damasio, *Looking for spinoza: Joy, sorrow and the feeling brain*. Orlando, FL, USA: Harcourt, 2003.
[58] J. Wang, M. Lewis, and S. Hughes, "Validating USARSim for use in HRI research," in *Proc. Hum. Factors Ergon. Soc.*, 2005, pp. 457–461.

Alexandre da Silva Simões (M'95) received the B.E. (1998) degree in electrical engineering from UNESP, Bauru, Brazil, and the M.Sc. (2000) and Ph.D. degrees (2006) in electrical engineering from the University of São Paulo (USP), São Paulo, Brazil.

He is a Full Professor with the Department of Control and Automation Engineering, UNESP. He is currently the Vice Director of the Sorocaba Campus of UNESP. His research interests include robotics, machine consciousness, machine learning, arts, and education. Dr. Simões was the General Chair of RoboCup 2014 and the founding president of RoboCup in Brazil.

Esther Luna Colombini received the M.Sc. (2005) and Ph.D. degrees (2014) in computing engineering from Technological Institute of Aeronautics, São Paulo, Brazil.

She is an Associate Professor with the University of Campinas (UNICAMP). Her research interests include machine learning, autonomous robotics, cognitive modeling, and robotics in education.

Dr. Colombini is the President of RoboCup Brazil and the Organizer of the Brazilian Robotics Olympiad, a public initiative that involves 100 000 students supported by the National Council of Technological and Scientific Development.

Carlos Henrique Costa Ribeiro received the M.Sc. (1993) degree in electrical engineering from Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, and the Ph.D. degree (1998) in electrical and electronic engineering from Imperial College London, London, U.K.

He is an Associate Professor with the Technological Institute of Aeronautics, São Paulo, Brazil. He has been a Researcher with the National Council of Technological and Scientific Development (CNPq), Brasìlia, Brazil, since 2001. His research interests include machine learning (especially algorithms for autonomous learning), multiagent robotics, complex networks, and engineering education.