International Journal of Machine Consciousness

A COMPUTATIONAL MODEL OF MACHINE CONSCIOUSNESS

JANUSZ A. STARZYK

School of Electrical Engineering and Computer Science,
Ohio University, Athens, OH 45701 USA
starzyk@bobcat.ent.ohiou.edu

DILIP K. PRASAD

School of Computer Engineering,
Nanyang Technological University, Singapore – 639798
dilipprasad@gmail.com

Received 15 October 2010
Revised Day Month Year

Despite many efforts, there are no computational models of consciousness that can be used to design conscious intelligent machines. This is mainly attributed to available definitions of consciousness being human centered, vague, and incomplete. Through a biological analysis of consciousness and concept of machine intelligence, we propose a physical definition of consciousness with the hope to model it in intelligent machines. We propose a computational model of consciousness driven by competing motivations, goals, and attention switching. We propose a concept of mental saccades that is useful for explaining the attention switching and focusing mechanism from computational perspective. Finally, we compare our model with other computational models of consciousness.

Keywords: Machine Consciousness; Computational model.

1 Introduction

Understanding consciousness and implementing it in manmade machines has interested researchers for a long time. However, since last two decades, research towards making machines conscious has gained great momentum and a lot of work is being done towards this end by various researchers [Anderson & Oates, 2007; Baars, 1998; Blackmore, 2002; Chrisley, 2003; Clowes, *et al.*, 2007; Dennett, 1993, 2001; Densmore & Dennett, 1999; Gamez, 2008; Haikonen, 2007; Rolls, 2007; Rosenthal, 1991; Sloman & Chrisley, 2003; Sun, 1997; Taylor, 2007; Velmans, 2002, 2009]. Despite significant efforts, research regarding computational modeling of machine consciousness is very limited. Among various technical, philosophical and computational difficulties, the primary reason is the difficulty in understanding consciousness and related abstract notions like thought, attention, awareness, etc. This paper addresses some of these issues as related to computational models of intelligence, and related modeling requirements for consciousness, attention and awareness.

At the outset, we recall the conclusion of the Schwarz Foundation meeting in 2001 ["The Schwarz Foundation: Can a machine be conscious?," 2001]: “The only (near) universal consensus at the workshop was that, in principle, one day computers or robots could be conscious. In other words, that we know of no fundamental law or principle operating in this universe that forbids the existence of subjective feelings in artifacts designed or evolved by humans."

Many researchers (philosophers, cognitive neuroscientists, psychologists, pharmacologist, physicist, artificial intelligence and cognitive scientist, etc.) tried to define or characterize consciousness in functional terms [Baars, 1998; Blackmore, 2002; Dennett, 2001; Sun, 1997; Velmans, 2002, 2009]. Many have discussed consciousness as causal (or non-causal) [Velmans, 2002], accessible (or inaccessible) [Dennett, 2001; Velmans, 2009], stateless (or having physical state) [Dennett, 2001; Velmans, 2009], representational (or non-representational) [Chrisley, 2003; Haikonen, 2007] and so on. However, nobody provided a complete theory of consciousness that may become a foundation for building computational models of conscious machines. Specifically, there is a lack of a physical description of consciousness that could become a foundation for such models. In our opinion, like intelligence, consciousness is a property of a physical mind not a meta-physical phenomenon. There are several reasons why a definition of consciousness will be useful, such as follows:

- To introduce a unifying definition of consciousness with minimum behavioral, design, and structural requirements that can be used as test of consciousness.
- To move from meta-physical to physical description of consciousness and obtain its computational model.
- To formulate a constructive approach and describe underlying mechanism that can result in machine consciousness.
- To describe consciousness as emerging phenomenon in the process of perception, learning, and building associative memories.

While the researchers may vary greatly in their opinions about machine consciousness and its feasibility, most of them identify the need for functional decomposition of the envelope term "machine consciousness" [Aleksander, 2009; Aleksander & Dunmall, 2003; Haikonen, 2000; Kuipers, 2008; Morsella, 2005; Sloman & Chrisley, 2003; Sun, 2007]. It is considered that various functional blocks (like embodiment [Chella, 2009], attention [Hudlicka, 2009; Taylor, 2007], emotion [Browne & Hussey, 2009; Chella, 2009; Haikonen, 2000; Hudlicka, 2009], experience [Albus, 2010; Chella, 2009], etc.) necessary for consciousness should be identified. Five axioms were proposed in [Aleksander & Dunmall, 2003] as essential functional properties of a conscious system. Requirements stated in these five axioms are roughly equivalent to embodiment and situatedness, episodic memory, attention, goals and motivation, and emotions. While the above functional properties are necessary conditions for a machine to be conscious, they do not form a sufficient set that is capable of imparting consciousness to a machine. Modular architectures were proposed in [Haikonen, 2000; Sloman & Chrisley, 2003; Sun, 2007; Taylor, 2007]. The architectures presented in [Haikonen, 2000; Sun, 2007; Taylor, 2007] focus upon identifying hierarchical levels in which various functional modules are distributed. On the other hand, Sloman and Chrisley [Sloman & Chrisley, 2003] concentrate on the nature of information that flows within and across various levels, while refraining from assigning specific functional modules to a particular level. In our opinion, conscious systems cannot be explicitly modeled as exclusively hierarchical. Every module must be connected to other modules with bidirectional flow of information, with a significant lateral as well as hierarchical flow.

We present our view of consciousness in relation to embodied intelligence ability to build stable sensory representations, and predict results of its actions in the environment. Self-organizing mechanism of emerging motivations and other signals competing for attention will be used to design models of conscious machines. In the next two sections, we briefly discuss the scientific and philosophical view of consciousness. The discussion presented is not exhaustive. It is rather representative of some important works that served as our inspiration. In the subsequent section, a discussion on emergence of consciousness is presented. This discussion helps us to understand consciousness from biological perspective. The section after this takes up the main task of defining consciousness. In this section, we first identify requirements for a conscious machine. Then, we build upon these requirements to define consciousness in physical terms. After this, we propose a computational model of machine consciousness based on our definition. This is followed by description of attention switching mechanism and mental saccades that drive sequences of conscious thoughts. Subsequently, we compare proposed model with other computational models of consciousness. Finally we conclude the paper.

2 Scientific view of consciousness

John Searle [33] said that ”Studying the brain without studying consciousness would be like studying the stomach without studying digestion, or studying genetics without studying the inheritance of traits”.

Marvin Minsky discusses consciousness in his book “The Emotion Machine” [Minsky, 2006]. He analyzes consciousness from the point of view of common sense, as well as presents views of other thinkers, philosophers, neuropsychologists, researchers in artificial intelligence and cognitive science. Their views of consciousness are different, from everything that makes human spiritual experiences, mystical links between sensing and highest levels of mind, to statements that “nobody has a slightest idea of how anything material can be conscious” [Fodor, 1992]. According to William Calvin and George Ojeman [Calvin & Ojemann, 1994], consciousness refers to focusing attention, mental rehearsal, thinking, decision making, awareness, alerted state of mind, voluntary actions and subliminal priming, concept of self and internal talk. Sloman [Sloman, 1991] suggests that it may be pointless trying to define consciousness, its evolution or function as they may have many different interpretations, similar to other big words like perception, learning, knowledge, attention, etc. Minsky points out that philosophers do not help in understanding consciousness, nor give recipe on how to test one.

Jeff Hawkins [Hawkins & Blakeslee, 2004] suggests that consciousness is a combination of self awareness and qualia (feelings associated with sensations but not related to sensory input). He also points out that consciousness is associated with declarative memory; the moment this memory is erased our conscious experience disappears. He suggests that memory and prediction play crucial roles in creating consciousness, whatever way one defines it.

Susan Greenfield's [Greenfield, 2000] concept, 'continuum of consciousness,' says that "Consciousness is a dynamic process and it changes with development of brain. Further, at macro-level there is no consciousness centre and at micro-level there are no committed neurons or genes dedicated to consciousness."

3 Philosophical view of Consciousness and Awareness

Higher-order theory supported by Rosenthal [Rosenthal, 1991] postulates the existence of a pair of distinct mental states: a first-order quasi-perceptual state, and a higher-order thought or perception representing the presence of that first-order state. In higher order theory "phenomenally conscious states are those states that possess fine-grained intentional contents of which the subject is aware, being the target or potential target of some sort of higher-order representation" [Rosenthal, 1991].

Baars [Baars, 1998] says that "Consciousness is accomplished by a distributed society of specialists that is equipped with working memory, called a global workspace, whose contents can be broadcast to the system as a whole". Further Baars says [Baars, 2002] that "only one consistent content can be dominant at any given moment." The content of the memory is decided by the consciousness. Dennett [Dennett, 1993] suggests that there is no single central place where conscious experience occurs; instead there are "various events of content-fixation occurring in various places at various times in the brain". When "content-fixation" takes place in one of these places, its effects may propagate, so that it leads to the utterance of one of the sentences that make up the story in which the central character is one's "self". Dennett believes that consciousness is a serial account for the brain's underlying parallelism.

Until now we had ignored an important question: Are consciousness and awareness the same thing? This question is important because many researchers often confuse these terms and use them interchangeably. In order to differentiate between consciousness and awareness, let's explore them from philosophical perspective.

Though most people perceive these two words as meaning basically the same thing, philosopher Nisargadatta [Nisargadatta, 1973] points to two very different meanings of consciousness and awareness. When he uses the term "consciousness", he seems to equate that term with the knowledge of "I Am". On the other hand, when he talks about "awareness", he points to the absolute, something altogether beyond consciousness, which exists non-dualistically irrespective of the presence or absence of consciousness. Thus, according to him, awareness comes first and it exists always. Consciousness can appear and disappear, but awareness always remains.

Our interpretation of his theory is that, contrary to the general belief, awareness is not a part (subset) of consciousness; in fact, he suggests that awareness is the superset of consciousness. In some sense he is relating awareness to something similar to a central executive, which remains whether we are conscious or not, and takes care of all the biological activities without actually making us conscious of them happening. We adopt his philosophy in our model.

Another important property of a conscious system in our model is intelligence. A system that is not intelligent cannot be conscious. A plant may be aware of a damage done to it, but it cannot be conscious about it, since it is not intelligent. In a similar way a worm cut in half is aware and may even feel the pain but it is not conscious about its own body being destroyed. Following this argument, in our model, a conscious system has to be aware as well as intelligent, but it need not be alive when it is implemented through a computational media in embodied machine, as schematically illustrated in Figure 1.

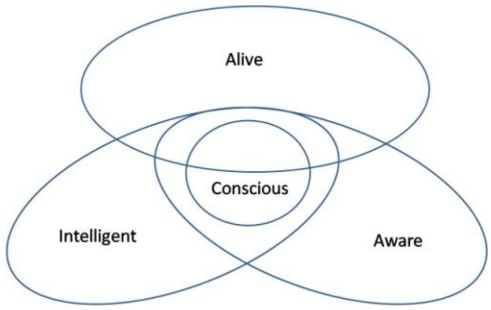

Figure 1: Consciousness centered view of intelligent systems

4 Emergence of consciousness

To further understand consciousness, we also need to understand - when does human consciousness emerge? What makes conscious beings aware of the fact that they are conscious? In the following, we discuss these questions from two viewpoints: developmental stages of human fetus and evolution of brain.

In human fetus, the first simple movement can be interpreted as the first indication of consciousness. Lou [Lou, 1982] argues against it. He argues that 8 weeks fetus' nervous system is in nascent state and its movement is the result of simple reflexes similar to the movement of headless chicken.

Development stages of brain indicate that consciousness can be a feature of cortex. In human fetus, cortex develops over many months and in stages. Spreen *et. al.* [Spreen, *et al.*, 1995] say that cortical cells come at the correct position in the 6^{th} week after gestation. 20^{th} week onwards, cortical region is insulated with myelin sheath and from 25^{th} week, the development of local connections between neurons takes place. 30^{th} week onwards, fetus' brain generates electrical wave patterns. These developments are gradual and incomplete until after birth. This makes it difficult to determine the exact moment of emergence of consciousness. On the other hand, we might conclude that this development may be the reason for limited consciousness exhibited by the fetus, and that consciousness emerges gradually with the development of brain.

Our analysis of relation between consciousness and the evolution of brain (based on [Greenfield, 2000; Lynch & Granger, 2008]) is summarized in Table 1. In this table, we also indicate the current ability of machines to model and mimic related evolutionary traits. This study is in agreement with Gerry Edelman [Edelman, 1992] who suggested that lower animals are not conscious. However, there are some interesting exceptions. For example, Young [Young, 1978] observed that octopus, though an invertebrate, possesses sophisticated learning and memory skills thus it may be considered as intelligent being. We also observe that most animals, including the animals with simpler brains, such as insects, exhibit some kind of circadian sleep wake cycle [Kaiser & Steiner-Kaiser, 1983]. This, in our opinion, is not sufficient to treat them as conscious beings according to our definition of consciousness.

Table 1. Evolution and consciousness.

		Living Being[1]	Evolutionary traits	Analogous feasibility in machines
↑	Conscious	Human Beings	Fully developed cross-modal representation Sensory capabilities: auditory, taste, touch, vision, etc. Bi-frontal cortex: planning, thought, motivation	Impossible at present
		Hedgehog (earliest mammals)	Cross-modal representation Sensory capabilities: auditory, touch, vision (less developed), etc. Small frontal cortex	Impossible at present
		Birds	Primitive cross-modal representation Sensory capabilities: auditory, touch, vision, olfactory. Primitive associative memory	Associative memories
	*	Reptiles [2]	Olfactory system Primitive vision	Computer vision (nascent)
	Not Conscious	Hagfish (early vertebrate)	Primitive olfactory system Primitive nervous system	Artificial neural networks
		Lower level animals (hydra, sponge, etc.)	Sensory motor units Point to point nervous system	Mechanical and/or electronic control systems

[1] Kingdom Animalia ; [2]* inconclusive\consciousness in transition

From the table, we conclude that some important traits necessary for consciousness include the presence of cortex, cross-modal representation, associative memory, and learning units.

Harth [Harth, 1993] has related consciousness to connections between the lateral geniculate nucleus (LGN) of the thalamus and the corresponding visual cortex. Thus we can conclude that the presence of cortex is important for emergence of consciousness.

5 Defining Consciousness

For designing models of conscious intelligent machines, we need to define consciousness in the context of such machines. We will adopt an approach similar to the one we took providing definition of intelligence [Starzyk, 2008], where our aim was not to remove ambiguity from philosophers' discussion about intelligence or various types of intelligence, but to describe mechanisms and the minimum requirements for the machine to be considered intelligent. In a similar effort, we will try to define machine consciousness in functional terms, such that once a machine satisfies this definition, it is conscious, disregarding the level or form of consciousness it may possess. Consciousness will be very much a function of embodied form of intelligence that a machine will possess. It will be an emerging property of machine development and its final form will depend on perception, memory, motor skills, motivations, thoughts, plans, etc.

In our prior work [Starzyk, 2008] on motivation in machine learning, we defined embodied intelligence as a mechanism that learns how to survive in a hostile environment. Thus, learning ability is a critical aspect of intelligent machine and knowledge is a result of learning.

5.1 *Requirements*

Since our aim is to design conscious machines, which exist and interact in an environment, we will use the most successful paradigm of building intelligent machines based on embodiment. Embodied intelligence uses sensors and actuators within its embodiment to perceive and act on the environment. In this interaction with environment, it learns to recognize objects and learns effects of its actions. By learning limitations of its embodiment and predicting how its own embodiment may affect the environment around it, machine learns to be aware of itself in a given environment. Thus in order to be aware of "who am I?" to predict the results of its own actions, to anticipate, etc., a *mechanism to acquire and represent knowledge* about the environment is required.

In order to be aware of "who am I?" attention to self is required. Similarly, for being conscious about other events/objects, attention needs to be given to those events/objects. Thus, *a mechanism for attention and attention switching* is required.

There are few interesting questions regarding attention and attention switching. When we are conscious of an object, our attention is focused on that object. During a thought process our attention switches from one object/fact/event to another one. What happens during this attention switch or what is the mechanism for attention switching during a thought process? Are we conscious during the attention switch? This should not be confused with D. Rosenthal's statement that "higher-order thoughts are themselves seldom conscious; so we are typically unaware of them" [Rosenthal, 1991] as he focuses on a conscious thought, not the underlying mechanism which creates a conscious thought. In our model there is no unconscious thought.

One possible explanation of the *mechanism of attention switching* is that it is a result of competition between *functional units* of brain. Attention switches to the semantic relation (object/fact/event) corresponding to the unit that wins. Sometimes, we think of an object/fact/event, but our attention drifts to something else, which was not directly related to this particular thought. This can be explained by the competition between numerous semantic relations as our *brain explores in parallel* how to reach a target goal or externally driven unconscious stimuli that switch our attention. We present our views on the state of consciousness during attention and attention switching in section 5.2 . For this purpose, we shall make use of the definitions of attention and attention switching.

Another requirement for consciousness is an act of cognitive perception, related to its actions, plans, expectations, and/or other cognitive experiences like thoughts or dreams. Cognitive perception or other cognitive experiences are related to *semantic memories* resulting from *knowledge building* and *episodic memories* of personal history related to past cognitive perceptions. In this model, *associative memory* is required to relate perception to knowledge.

Even as we discuss various requirements for consciousness, we need to note that brain can't be easily compartmentalized, i.e., there is no point-to-point connection between senses and neurons in brain. For example, visual system activates at least 30 different areas of brain. Similarly, no single brain region (in agreement with Susan Greenfield [Greenfield, 2000] and D. C. Dennett [Dennett, 1993]) or no single chemical process in the brain is responsible for consciousness. We identify that consciousness is a result of interactions between interconnected modules with attention switching mechanism helping to select a cognitive experience and managing a sequential cognitive process. This point confirms with the Dennett's believes of consciousness to be a serial manifestation of brain's parallelism [Dennett, 1993].

5.2 *Attention and attention switching*

Attention is a selective process of cognitive perception, decision making, action control or other cognitive experiences. This selective process of attention results from attention switching (needed to have a cognitive experience).

Attention switching is a dynamic process resulting from competition between representations related to motivations, sensory inputs and internal thoughts including spurious signals (like noise or unexpected external stimuli). Thus attention switching may be a result of deliberate cognitive experience (and thus fully conscious signal based on knowledge, prediction, association or episodic memory) or it may result from subconscious process (stimulated by internal or external signals). Thus, while paying attention is a conscious experience, switching attention does not have to be.

Note, that in this respect the major mechanism responsible for conscious thoughts and their dynamics combines both top-down (conscious) and bottom-up (unconscious) signals. This is different from Rosenthal's HOT theory [Rosenthal, 1991] or Baar's

Global Workspace theory [Baars, 2002] which are entirely driven by conscious thoughts. Our approach is closer to Dennett's "frame in the brain" idea [Dennett, 1993] where coalitions of neurons compete for the frame with winners becoming a conscious experience.

5.3 *Who am I?*

As discussed, to be conscious, the ability to define "Who am I?" is essential. In functional terms, it means that the system can perceive, act, and predict results of its actions. This includes the effect of its own embodiment on the environment and limitations resulting from constraints of this embodiment. It also contains predictions of how own embodiment is perceived by other individuals in this environment. A large animal may easily predict reaction of smaller animals that it encounters and becomes aware of its size, physical strength, or fearsome posture; indirectly defining "Who am I?" and developing self awareness.

Thus, self-awareness is a synonym for consciousness in its minimalistic functional meaning. Following Nisargadatta, we accept general awareness as a prerequisite for consciousness. However, we note that although a plant is aware of light and cold, yet it is not conscious. Consequently, in our definition, consciousness requires cognitive awareness.

5.4 *Central executive*

A central executive, which operates no matter whether we are conscious or not, is required as the platform for the emergence, control, and manifestation of consciousness. In human, central executive takes care of all the biological activities without making us aware of what is happening, as well as of all cognitive perceptions, thoughts and plans. In machine, central executive will control its conscious and subconscious processes, driven by its learning mechanism, creation and selection of motivations and goals. Thus, central executive, using cognitive perception and cognitive understanding of motivations, thoughts or plans is responsible for self-awareness and creates conscious state of mind.

5.5 *Definition of Consciousness*

We define machine consciousness as follows:

A machine is 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; the central executive is driven by the machine's motivation and goal selection, attention switching, semantic and episodic memory and uses cognitive perception and cognitive understanding of motivations, thoughts, or plans to control learning, attention, motivations, and monitor actions.

Thus, central executive, by relating cognitive experience to internal motivations and plans, creates self-awareness and conscious state of mind.

6 Computational model of consciousness

In this section, we propose a computational model that integrates functionalities of biological systems in a virtual machine. We are taking functional inspiration from biological systems to model consciousness and realize it in machines. It should be noted that we are not advocating mimicking biological systems exactly in machines, but rather to model and use their functional organization of conscious processing.

Our model consists of three main functional blocks, viz., Sensory-motor, Episodic Memory and Learning, and Central executive. A detailed block diagram of the model is presented in Figure 2.

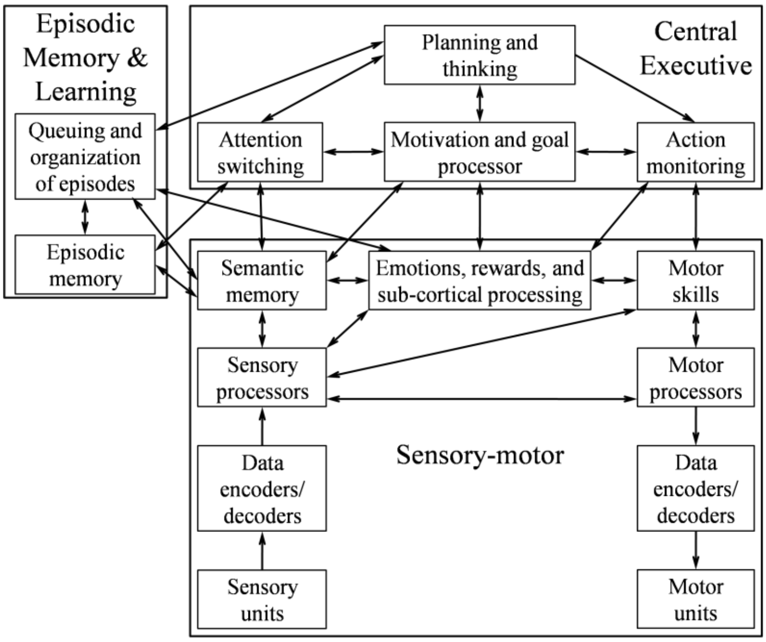

Figure 2: Computational model of a conscious machine.

6.1 *Sensory-motor block*

Sensory-motor block is composed of three parts: sensory processors integrated with semantic memory, motor processors integrated with motor skills, and sub-cortical processor integrated with emotions and rewards.

The sensory processors are connected to the sensors through encoders/decoders. They receive the sensory data and are responsible for concept formation in self-organized hierarchical structure [Hawkins & Blakeslee, 2004]. Sensory processors build and activate associative semantic memory that represents knowledge about the environment. For achieving this, sensory processors interact with various units in central executive (like attention switching, motivations, and goal processor units), episodic memory, and sub-cortical processors (responsible for emotions and rewards). Semantic memory blocks activated by various sensory inputs are interconnected with each other making the cross-modal representation in such systems possible.

Similarly, the motor processors are connected with the motor units through encoders/decoders. The processors actuate the motors and receive feedback through sensory processors. Through interaction with central executive and sub-cortical processor they build a hierarchy of motor skills. Cognitive action monitoring is supervised by central executive.

The processors in the emotions, rewards, and sub-cortical processing are used for generation of emotional and reward signals that govern learning and serve as an interface to other units. Specifically they cue episodic and semantic memories, switch attention, provide motivations, help to create and select goals, and interact with action monitoring.

Multiple processors in the sub-cortical processor block execute their programs in parallel and generate individual outputs. These outputs may compete among themselves (at a subconscious level) or may be used by the central executive unit (at a conscious level) to make a selection. Such a mechanism shall be helpful for attention, goal selection, motivation, etc.

6.2 *Episodic memory and learning block*

Episodic memory and learning block is composed of two parts, episodic memory unit and learning unit. Episodic memory unit is a collection of smaller functional blocks, each dedicated to capture a spatio-temporal sequence of semantic relationships like relations between objects observed in an episodic experience with their significance derived from emotional context.

Cueing and organization of episodes unit is able to recognize the novel events/patterns in various processes and help to build semantic relationships. For doing so, it shall collect and process data from all the units including perceptions, motivations and interpretations of cognitive experiences and their significance from the central executive. Subsequently, it stores and manages episodes and also initiates learning about specific events/patterns if directed by the central executive.

6.3 *Central executive block*

Central executive block is responsible for coordination and selective control of all the other units. This block interacts with other units for performing its tasks, gathering data and giving directions to other units. Its tasks include cognitive perception, focusing attention, attention switching, motivation (based on internal goals [Starzyk, 2008] and winner-take-all mechanism), goal creation and selection, thoughts, planning, learning, etc. For this purpose, it needs the capability to dynamically select and direct execution of programs that govern attention, motivation, episodic memory and action monitoring. In addition, central executive can activate semantic memory and control emotions.

Central executive directs cognitive aspects of machine experiences but its operation is influenced by competing signals representing motivations, desires, and attention switching that are not necessarily cognitive or consciously realized. Central executive does not have any clearly identified decision making center. Instead, its decisions are result of competition between signals that represent motivations, pains and desires. At any moment, competition between these signals can be interrupted by attention switching signal. All such signals constantly vary in intensity as a result of internal stimuli (e.g., increasing hunger) or externally presented and observed opportunities (fear, lucky chance). Thus, the fundamental mechanism that directs machine in its action is physically distributed as competing signals are generated in various parts of machine's mind. Further, it is not fully cognitive, since, before a winner is selected, machine does not interpret the meaning of competing signals.

Cognitive aspect of the central executive mechanism is predominantly sequential, as a winner of the internal competition is identified and serves as an instantaneous *director* of the cognitive thought process, before it is replaced by another winner through attention switching and changing motivations to act.

Once a winner of internal competition is established, central executive provides cognitive interpretation of the result, providing top down activation for perception, planning, internal thought or motor functions. It is this cognitive realization of internal processes that results in central executive's *decision* of what is observed, *planning* how to respond, internal *talk* of what to do, that we associate with a conscious experience and a continuous train of such experiences constitutes consciousness. Next section provides detailed discussion of the mechanism that produces this continuous train of conscious experiences.

It should be noted that though the sensory and motor units are not the part of the brain as defined above, they are essential for embodiment of the complete system.

7 Attention switching mechanism of the proposed model

Focus of attention and ability to switch attention are naturally associated with working of a conscious mind. There are various hypotheses of how to model attention switching and how anything may be selected as a focus of attention [Albus, 2010; Baars & Franklin, 2007; Desimone & Duncan, 1995; Gamez, 2008; Posner, 1994; Taylor, 1997, 2007]. Taylor [Taylor, 1997] models it in the form of control structures that can perform two things: first, they can produce global correlations between various working (semantic) memories; second, they can support competition between various inputs such that the memories most relevant to the winner of competition are activated. The first part provides the function of attention focusing and the second part provides the function of attention switching. In the CODAM model proposed by Taylor [Taylor, 2007, 2009], the attention focusing and switching is accomplished using neural circuits that can simulate a corollary discharge module capable of realizing both the functions (details in section 8.5). Albus [Albus, 2010], on the other hand relies strongly on the biological model of attention focusing based on the lateral inhibition of the neurons. Apart from the lateral inhibition model of visual attention at the retina, there are mechanisms inside the brain that are structured to focus the attention. The thalamus plays the central role in focusing the attention upon an important goal by selectively suppressing other unimportant sensory signals on their way to the cortex.

In this section we illustrate how the proposed computational model of consciousness can implement both attention focusing and attention switching.

7.1 *Attention switching signals*

Attention switching is accomplished as a result of competition between various types of signals, viz., internal thought signals, new perceptions due to external environmental changes, and motivational changes.

First and foremost, attention switching is effected by internal thought signals that come from cognitively selected spotlights of attention. They are normally associated with conceptual problem solving, where various issues are cognitively represented and manipulated in relation to the selected goals and internal motivations that drive the thinking process. They correspond to mental searches for the most useful approach to the problem. They are frequently changed as various alternatives are considered and their relevance to the selected problem is evaluated.

The second type of signals that compete for attention switching are new perceptions. New perceptions represent the observed changes in the environment and they typically happen as results of actions taken by the system acting on the environment. New perceptions are less frequent than changes in the inner thoughts and while some of them may be represented as new cognitive attention spots, others will enter competition without being cognitively recognized. Due to these signals, attention switching may be a result of sudden changes in the external environment that indicate possible threat or unexpected chance. These signals, once recognized cognitively, may change the motivation or action selected by the system.

Finally, the third type of signals that compete for machine attention and may result in attention switch are changing motivations. Motivations may be change due to external or internal conditions that affect the system of values or due to cognitive modulation. Motivations are related to either cognitively recognized or unconscious needs and desires, emotional changes, and internally set abstract goals.

Figure 3 illustrates the three types of attention switching signals. As a result of attention switching, system focuses its cognitive attention on the selected issue (conceptual thought, observation, or need). A result of cognitive attention may be a decision to act that stimulates schemas for action and action supervision. In addition to feedback from cognitive attention, motivations and perceptions are influenced by external and internal signals.

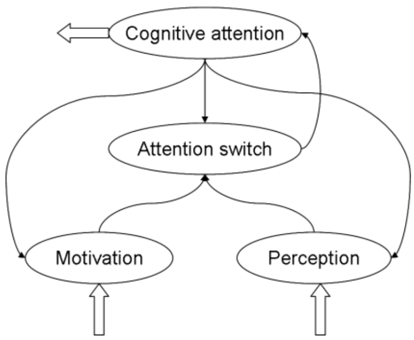

Figure 3: Attention switching signals

7.2 *Mental saccades*

Presented in this section is the mechanism of mental saccades that uses the idea of global workspace introduced by Baars [Baars, 1997; Baddeley, 1993]. Saccadic eye movements are well researched aspects of selective visual perception. They result in scanning a visual scene providing focus for visual inspection. Saccades are also used to study motor control, cognition and memory in conjunction with functional imaging and transcranial magnetic stimulation [Leigh & Kennard, 2004]. Eye movements reflect a large number of psychological processes underlying various cognitive tasks. For instance, in [Liversedge & Findlay, 2000], it was observed that saccadic eye movement data are a good indication of cognitive process during visual search and reading. It was determined that "where" and "when" aspects of saccades are separate both psychologically and physiologically and depend on attentional selectivity [Desimone & Duncan, 1995] and semantic properties of the fixated objects [De Graef, *et al.*, 1990]. Saccades help determining cognitive processing time and relationship between attention and eye movements [Liversedge & Findlay, 2000].

We hypothesize that a similar search that drives motor control of the eye movement takes place in the mental workspace. We define this scanning of the mental workspace as mental saccades. In our model, mental saccades constitute a core mechanism for conscious selection of attention spotlight and drive the machine consciousness.

To understand mental saccades, let us consider Figure 4 that illustrates the activation of memory traces in the frontal cortex area (global workspace) based on the observed scene. An input image is shown in Figure 4 together with selected part of this image (a male figure) that corresponds to the visual attention resulting from an eye saccade. Perceived input activates corresponding area of the face recognition and associated areas of semantic and episodic memory that relate to the recognized person. This in turn activates memory traces in the global workspace area that will be used for mental searches (mental saccades).

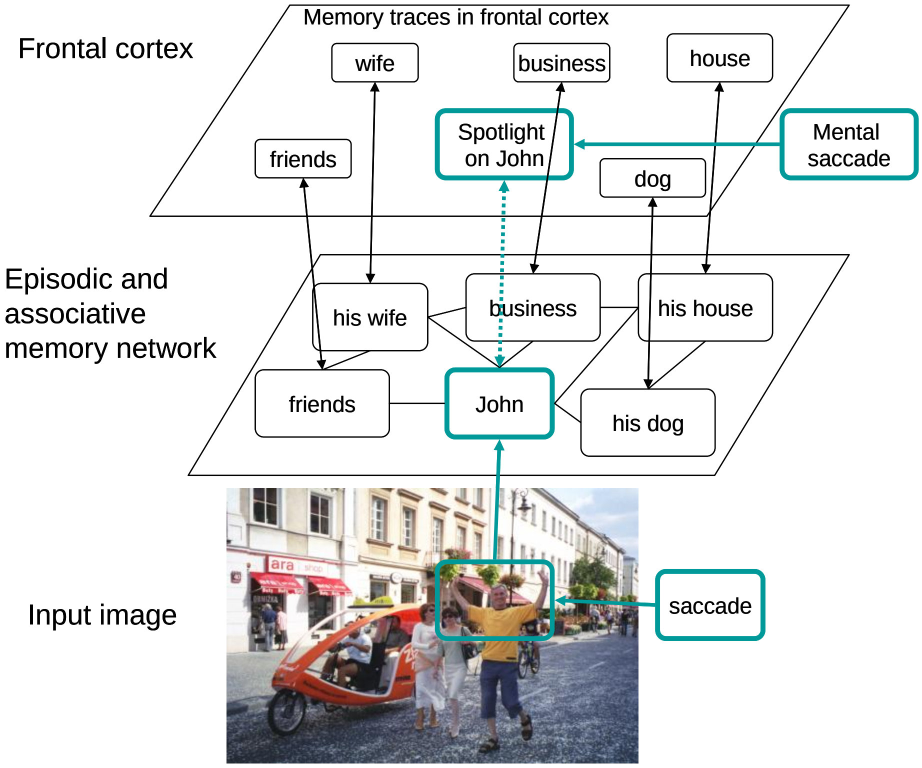

Figure 4: Selection of the cognitive spotlight

Activated memory traces in the global workspace area are searched in a process similar to visual saccades. For instance, in the illustrated activation in Figure 4, a mental saccade may proceed from the observed person (John) to one of the concepts associated with this person, e.g. his dog. Thus a mental spotlight will move from John to his dog and its relation to the current objective will be evaluated. This in turn may activate different semantic and episodic memories related to the dog through a feedback from the memory trace. To avoid returning to the same concepts in the working memory, previously scanned memory traces will be inhibited until the working memory is searched through.

Mental saccades constitute an alternative solution to the idea of attention codelets competing for consciousness as presented in the LIDA model [Baars & Franklin, 2003]. What differentiates our approach from the LIDA model is that we use explicit mechanism that switches machine attention to representations that can activate memory traces in the global workspace. In particular, we do not need to refer to ill-defined cognitive "atoms" or "attention codelets" needed to begin the consciousness phase in LIDA, nor do we need to refer to vague terms like "contents of consciousness" to describe the conscious state of mind. In our model there is no movement of selected portions of current situational model to the global workspace required in LIDA model. Situational model (as discussed in LIDA) or its portion would be difficult to isolate from complex interactions between motivations, observations and inner thoughts.

We prefer the saccading search rather than competition, since it is not clear how unconscious perceptions may compete with each other. However, the two theories are close in the sense of bringing a selected cognitive experience into attention spotlight and by doing so making it a conscious experience.

Next we discuss the organization of cognitive process based on mental saccades and attention switching mechanisms.

7.3 *Cognitive process*

Figure 5 illustrates major associations that take place in the use of mental saccades as a driving mechanism for attention focus and conscious mental process. While we do not posit that mental saccades can be observed in the human brain, the proposed mechanism is computational and explains how a conscious process can direct machine observations, motivations and motor functions.

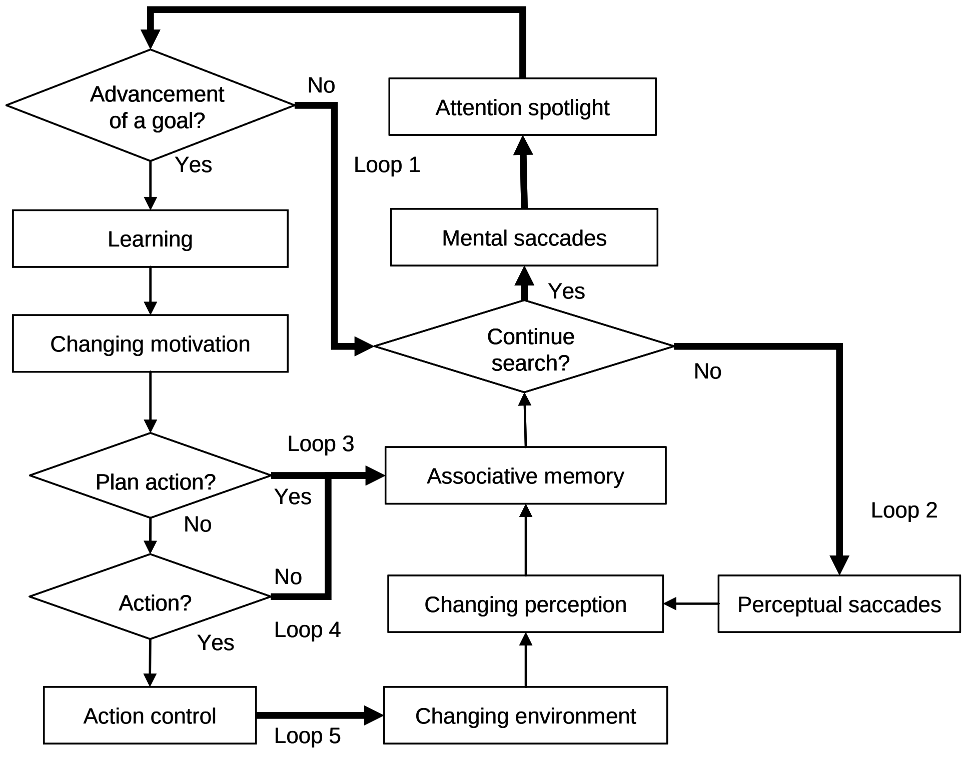

Figure 5: Mental saccades in a conscious machine model

In what follows we will trace major loops in Figure 5 to explain possible outcomes of a cognitive cycle and dynamic changes that may result from changes in the environment, internal motivations and mental probing of alternative solutions to a selected problem.

1. Mental saccades constitute the innermost, most frequently updated loop of attention focus (Loop 1 in Figure 5). They activate a selected memory trace and bring it to a conscious attention spotlight. Attention spotlight (object, action, event, idea) is then compared to the current goal and its ability to advance the goal is evaluated. In a typical saccadic mental search, the search continues until a goal is advanced by one of the cognitively selected ideas or a change in the mental scene is observed. Possible advancement of a goal is followed by mental reinforcement and learning of what is expected to advance the given goal. Possibly a new goal is mentally set in the block called "changing motivations". Notice, that "changing motivations" can also be updated independently from the mental process by changes in the environment or internal state of the machine. This inner most "mental" loop that compares various alternative solutions corresponds to action planning and does not have to lead to the execution of an action by activation of motor control functions. For instance, it can lead to a "theoretical solution" that is conceptualized and perhaps stored in the episodic memory.
2. Loop 2 is associated with changing perceptions due to saccading movement. Such changes may occur either voluntarily, when conscious examination of a given mental scene is completed or as a result of attention switch (conscious or subconscious).
3. If the idea in attention spotlight was deemed to advance a current goal, then a decision is made whether this idea leads to a multistep solution and requires planning for the next step. If yes, then each required step is analyzed by bringing its elements into attention spotlight one by one within the third loop (Loop 3). A cycle in Loop 3 is completed by bringing new scenario into associative memory which starts a new series of mental saccades related to this new step in conscious search for a solution.
4. If the action plan is completed, then a decision is made whether the obtained solution to the current goal should be implemented by performing an action. If not, then the machine completes Loop 4 with possible update of sequential memory resulting from learning the results associated with mental task completed in this loop.
5. Loop 5 is completed if a decision to start an action is reached after mental evaluation. Acting on the environment results in changing perception, which updates associative memory contents and forms the basis for new mental saccades.

Besides the described typical flow of the conscious process, changes in the environment that result from execution of a task or external interference automatically influence the machine's perception and indirectly its conscious process. In fact, if it is desired that a machine focuses on a mental task that requires isolation from ongoing changes around it, a separate blocking mechanism must be built to protect the machine from such unwanted interruptions. In addition, changes in the internal state of the machine that are critical element of the "motivated learning" [Starzyk, 2010] influence the cognitive cycles by changing either motivations or the content of the working memory. What is important is that between the changes in the environment and the internal state of the machine, one can be almost certain, that the cognitive process will never stop.

8 Comparison with other computational models

In this section, we compare the computational model of consciousness presented in this paper with axiomatic theory of Aleksander [Aleksander, 2009; Aleksander & Dunmall, 2003] and other computational models like CogAff schema [Sloman & Chrisley, 2003], Haikonen's model of consciousness [Haikonen, 2003], LIDA model of the global workspace theory [Baars & Franklin, 2009], and CODAM model [Taylor, 2009].

8.1 *Aleksander's axioms of machine consciousness*

Aleksander has contributed significantly to the development of an axiomatic theory of machine consciousness. Though he has not proposed a computational model of machine consciousness, he has proposed five axioms as the minimum conditions for machine consciousness [Aleksander, 2009; Aleksander & Dunmall, 2003]. Thus, we begin the comparison of our computational model of machine consciousness against this benchmark. Aleksander's five axioms [Aleksander, 2009; Aleksander & Dunmall, 2003] can be summarized as follows:

1. **Presence of the agent in the world**: This axiom is closely related to the embodiment and situatedness of the agent in the physical world.
2. **Imagination and recall of past and present experiences**: In our opinion, this axiom considers two aspects. First aspect is the storage of the events (and experiences of the conscious agent) in the form of associative memories and the ability to recall them when queried. Second aspect is the ability to formulate imaginary scenarios and predict the outcomes of the scenarios based on the experiences.
3. **Attention as a determiner of the content of experience**: This axiom requires that the agent interacts (using sensory data collection, motor actions, and/or sub-cortical processing) with the object or event of current attention spotlight.
4. **Volition**: The agent should have will, motivation, and capability to form goals. Further, the agent should be able to take certain actions to fulfill its goals.
5. **Emotion**: The agent should be able to identify the nature of experiences in its own context and set of emotions.

Here we elaborate how our model is in agreement with all the above axioms. Since our model considers embodiment as one of the minimum requirements for a conscious machine, the first axioms that requires situatedness is already considered in our model.

With regard to the second axiom, the requirement of associative memory is already addressed in sections 5.1 and 6.2 . The planning and thinking unit in the central executive block is capable of generating hypothetical scenarios required for imagination. The ability of the central executive block to dynamically select and direct execution of programs (that govern attention, motivation, episodic memory, and action monitoring), activate semantic memory, and control emotions together give it the ability to imagine.

The contents of the third axiom are already discussed in detail in section 5.2 and 7. Here, we mention succinctly that the attention focusing and switching mechanism stimulate the schemas for action and action supervision and hence act indirectly as the determiner of the content of experience.

The central executive block contains the functional units for motivation, goal creation, attention switching, planning, thinking, action monitoring, which together imparts volition to the agent.

Emotions are incorporated in our model in the sensory-motor block. The unit for emotions, rewards, and sub-cortical processing interacts directly with the semantic memory (which contains the semantic information about the object or event in current attention), motivations and goal processor unit, and action monitoring unit to generate emotional states of the agent.

8.2 *CogAff Schema [Sloman & Chrisley, 2003]*

Sloman and Chrisley [Sloman & Chrisley, 2003] presented a CogAff schema for conscious virtual machines. As discussed in [Sloman & Chrisley, 2003], CogAff is not an architecture or a functional model of consciousness. It is rather an ontological model of consciousness. On the other hand, our model is a functional model of consciousness. Thus, the CogAff schema and our model cannot be compared directly. However, if we study the ontological structure of our model, the ontological nature of the processes in the individual functional units, and the ontology of the information interchange between various functional blocks, our model can be represented ontologically in a model similar to the CogAff schema. Drawing an ontological analogy between our functional model and the CogAff architecture require detailed study of the ontology of our model and may constitute a future work.

8.3 *Haikonen's model of consciousness [Haikonen, 2003]*

Haikonen [Haikonen, 2003] says that consciousness is not implemented as a single functional model. It rather emerges when the cognitive processes of the agent work in unison and focus their attention upon the same entity as perceived by the various sensory modalities. While we agree that consciousness emerges due to the various processes acting in unison, in our model, self awareness and conscious state of mind emerge due to the central executive relating cognitive experience to internal motivations and plans.

Haikonen indentifies attention, emotion, and introspection as essential components for emergence of consciousness. Our model has functional blocks for attention and emotion. Though our model does not have an explicit functional unit for introspection, the units in the central executive are capable of realizing introspection by activating learnt knowledge of self in relation to perceptions, thoughts and planned actions. Thus, our model is in agreement with the general concept of consciousness described by Haikonen.

A significant difference between the two models is that Haikonen model assumes that the focus of attention is a result of competition between various sensory modalities and that the winner broadcasts its contents to other modules and "recruits" them to respond to his perception. In our model, competition is not between various sensory modalities but between various kinds of signals that includes inputs from sensory modules, but also uses competing motivating pain signals as well as current winners of cognitive focus of attention. Only the last signal is fully cognitive, while others are not consciously defined unless they become winners in this internal competition. This difference is significant since, in our model, goal driven motivations play decisive role for selecting the focus of attention most of the time. Sensory inputs provide many disrupting signals, that are used by the conscious system to evaluate the changes that happen in the environment either as a result of its own actions or unexpected changes caused by other actors. In fact, Haikonen states directly that "the consciousness … does not have causal power". We disagree with this statement, as a conscious mind actively considers various options in relation to goals and directs body to perform useful actions.

Haikonen's model has other problems as well, that cannot be easily solved. To allow for the execution of subconscious operation in parallel to a conscious one, his model assumes that "some modules engage in mutual communication while other modules are doing something else". This violates a single conscious focus of attention. While such partial cooperation between modules is possible in concurrent systems, it may lead to split personality as a single cooperation group may dominate in one instance and other in the next one. In our model, a single winner directs the focus of attention, and concurrent subconscious operations are used to control routine tasks that do not require central supervision and cognitive attention.

We agree with Haikonen that design of "some exotic qualia process" is not necessary and that sensed qualities that people associate with qualia can be presented directly to the system in the form of perceptional representations. We also agree with his analysis of free will as a result of conflicting interests and some degree of randomness, but basically realized by cognitive system as the "own will".

Our final disagreement is in his rejection of the need or usefulness to build machines that can feel pain, as our system uses pain based motivations for development and direct control of attention switching in the conscious machine.

8.4 *LIDA model [Baars & Franklin, 2009]*

Baars and Franklin [Baars & Franklin, 2009] presented the LIDA model as a computational model of consciousness based on the global workspace theory (GWT). A cognitive cycle in LIDA is subdivided into three phases: the understanding phase, the consciousness phase, and the action selection phase.

In the first phase, i.e. the understanding phase, the agent creates its perception of the sensory input using the associative memory. Then, the agent uses its episodic and declarative memory to create local associations in the workspace, which are used to generate the agent's understanding of the current situation (what is going on right now).

In the second phase, which the authors call the consciousness phase, attention codelets are used to pass coalitions of selected portions of the understanding of the current situation to the Global workspace. In the same phase, the coalitions compete in the global workspace and the coalition that is most relevant to the agent in the current situations is selected. This most relevant coalition and its contents are broadcast globally as the content of consciousness.

In the third stage, called the action selection phase, the conscious broadcast of consciousness (begun in the second phase) triggers the perceptual associative memory and begins the process of learning (new entities and associations) and reinforcement (of existing entities and associations). The transient episodic memory encodes the conscious broadcast as new memories and the procedural memory stores the new action schemes.

In their model, before getting the attention (Attention Codelets), which begin the consciousness phase, the agent already has an understanding of the current situation without needing consciousness or attention. In this model, the data from the workspace compete and then selection is done in the Global Workspace to decide the currently most relevant coalition. The selected coalition becomes the content of consciousness. Effectively, in their model, consciousness plays the role of only broadcasting the selected coalition in the Global Workspace.

The Global Workspace in the LIDA model may be considered partially analogous to the Central Executive in our model, because both are the decision making blocks in the respective models. However, besides this rough analogy and the presence of attention (though Attention Codelets in the LIDA model), the two models are totally different. In our model, the agent does not understand the current situation without attention. Further, attention is not the trigger to consciousness in our model. Rather, in our model, motivation to act and attention switching mechanism are responsible for selecting the focus of attention.

While LIDA model considers consciousness as a phase in the cognitive cycle of the agent, in our model it is rather an emergent state of mind. In the LIDA model, the contents of consciousness generated in the consciousness phase are used in the action selection phase for updating associative, transient episodic, and procedural memories [Baars & Franklin, 2009]. In our model, planning, updating of associative memories, and action selection are mainly driven by the motivations and goal planning unit and the attention switching mechanism. Thus, as compared to the LIDA model where consciousness is a phase in the cognitive cycle influencing the action selection phase, we consider consciousness as a state of agent which emerges by relating the cognitive experience to the internal motivations and plans, with reference to self.

We also differentiate our model from the LIDA model in section 7.2 where we discuss that our model of conscious attention focusing is based on the search of mental saccades, while the LIDA model uses competition of attention codelets. Though these two approaches towards modeling the attention are close to each other in theory, our model does not require complicated and ill-defined entities (such as 'attention codelets' or 'contents of consciousness') and does not involve partial complex representation movement of data across distributed memories.

8.5 *CODAM model*

In [Taylor, 2009], Taylor says that, "Consciousness arises through a speeding-up process by employment of a more powerful attention control than of purely ballistic form, using a corollary discharge of the movement of the focus of attention." Taylor draws a rough (though unproven) analogy of attention movement model with the nucleus reticularis thalami (NRT) cortex in the context of visual attention, in which lateral inhibition dampens any localized activity thus helping the brain to focus onto a global goal of attention focusing. It is the same lateral inhibition that controls the local competition of the signals, such that the most prominent signal activates a corollary discharge and emerges as the winner (which then becomes the global goal of attention focusing). In the context of the control system (for machine consciousness), he suggests that this function can be utilized using a corollary discharge module [Taylor, 2007, 2009]. In the CODAM model [Taylor, 2007, 2009], a corollary discharge unit is a unit that acts as a buffer of the input feedback signals (generated by an action monitoring module). These input signals act as error signals. As long as the net signal (equivalent to the result of the competition of various signals) is below a certain threshold, the whole agent keeps focusing on the current global goal of attention by error correction. In this sense, corollary discharge module acts as a predictor of the current goals. However, if the net signal exceeds the threshold value, a corollary discharge is initiated, which helps choosing the current winner among the competing signals and make it as the next global goal of attention focusing, thus realizing attention switching.

The CODAM model [Taylor, 2009] does recognize the importance of attention, attention control, goals, and emotions. Moreover, it suggests that attention may be sufficient for consciousness. In the context of our model, while these components are necessary for the emergence of consciousness, they are not the sufficient conditions. In our model, conscious mind generates focus of attention based on the interplay of motivations and attention switching that selects what to pay attention to. Thus attention is an expression of a conscious effort not its trigger.

Another reason we believe that CODAM oversimplifies the consciousness phenomenon is that "the ownership aspect occurs in association with a given input being attended to" that is particular to each stimulus. Thus activation of "self" is particular to each stimulus and the proposed in CODAM coding of "I" as context free is related to the "nothingness" [Sartre, 1943] and ultimately to the lack of content. As a result, Taylor cannot recommend either of the two solutions he considered to determine "self". In our model such problem does not exists as the "self" component of the attended stimuli comes after attention was focused on this specific stimuli in a top-down fashion.

In spite of these differences, we agree with Taylor when he states that consciousness is the topmost controller. Perhaps the differences in our models stem from how we approach the role of attention. In our model, we use two separate mechanisms related to attention: the first one is attention switching that is triggered by both conscious and unconscious signals (and thus it is controlled by conscious mind as well as external stimuli), the second one is focus of attention (subject solely to a conscious effort) which is a result of attention switching and "acts as a filter" for the lower cortical levels (as it is in CODAM).

9 Conclusion

Opposed to the interpretation of consciousness as a metaphysical phenomenon, we present its physical definition based upon the biological study of conscious brains and our model of embodied intelligence [Starzyk, 2008]. Though, our definition of consciousness parallels a biological perspective, the proposed definition is fully computational and clearly encompasses various phenomenological characteristics of consciousness.

Our proposed organization of conscious machine model is based on two important observations. First, biological evolution as well as development of human brain indicates that a functional unit similar to pre-frontal cortex is responsible for the emergence of consciousness. Thus, in our model, consciousness is an emerging phenomenon. Second, a central executive which controls and coordinates all processes, whether conscious or subconscious, and which can perform some of its tasks (like memory search) using concurrent dynamic programming, is necessary for developing consciousness. Such unit in our model uses distributed and competing signals that represent goals, motivations, emotions, and attention. Only after the winner of such competition is established, it drives a focus point of a conscious experience.The proposed computational model of consciousness mimics the biological systems functionally and retains a well-defined architecture necessary for implementing consciousness in machines. We compare our model with various existing models of consciousness in order to highlight the important and distinct features of our model. Our model might be neither complete or foolproof nor practically feasible. However, it should provide guidance towards building models of conscious embodied machines.

In addition to the computational model of consciousness, we propose a concept of mental saccades to explain the attention switching and focusing in our computational model. Our model uses the competition among three different types of signals in the cognitive cycle of the agent. However, the exact mechanism of attention switching is explained using mental saccades, which may not relate directly to human consciousness, but are useful for the computational implementation of consciousness in machines.

References

Albus, J. S. [2010]. A model of computation and representation in the brain. *Information Sciences, 180*(9), 1519-1554.

Aleksander, I. [2009]. The potential impact of machine consciousness in science and engineering. *International Journal of Machine Consciousness, 1*(1), 1-9.

Aleksander, I., and Dunmall, B. [2003]. Axioms and tests for the presence of minimal consciousness in agents. *Journal of Consciousness Studies, 10*(4-5).

Anderson, M. L., and Oates, T. [2007]. A review of recent research in metareasoning and metalearning. *AI Magazine, 28*(1), 7-16.

Baars, B. J. [1997]. *In the Theater of Consciousness: The Workspace of the Mind*. New York: Oxford University Press.

Baars, B. J. [1998]. *A cognitive theory of consciousness*: Cambridge University Press.

Baars, B. J. [2002]. The conscious access hypothesis: Origins and recent evidence. *Trends Cogn Science, 6*(1), 47–52.

Baars, B. J., and Franklin, S. [2003]. How conscious experience and working memory interact. *Trends in Cognitive Sciences, 7*(4), 166-172.

Baars, B. J., and Franklin, S. [2007]. An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA. *Neural Networks, 20*(9), 955-961.

Baars, B. J., and Franklin, S. [2009]. Consciousness is computational: The LIDA model of global worskspace theory. *International Journal of Machine Consciousness, 1*(1), 23-32.

Baddeley, A. D. [1993]. Working memory and conscious awareness. In A. Collins, S. Gathercole, M. A. Conway & P. E. Morris (Eds.), *Theories of Memory*. East Sussex, UK: Psychology Press Ltd.

Blackmore, S. [2002]. There is no stream of consciousness. *Journal of Consciousness Studies, 9*(5-6), 17-28.

Browne, W. N., and Hussey, R. J. [2009]. Emotional cognitive steos towards consciousness. *International Journal of Machine Consciousness, 1*(2), 203-211.

Calvin, W. H., and Ojemann, G. A. [1994]. *Conversation with Neil's brain: the neural nature of thought and language*: Addison-Wesley.

Chella, A. [2009]. Machine consciousness: a manifesto for robotics. *International Journal of Machine Consciousness, 1*(1), 33-51.

Chrisley, R. [2003]. Embodied artificial intelligence. [Review]. *Artificial Intelligence, 149*(1), 131-150.

Clowes, R., Torrance, S., and Chrisley, R. [2007]. Machine consciousness - Embodiment and imagination. [Editorial Material]. *Journal of Consciousness Studies, 14*(7), 7-14.

De Graef, P., Christiaens, D., and d'Ydewalle, G. [1990]. Perceptual effects of scene context on object identification. *Psychological Research, 52*, 317-329.

Dennett, D. C. [1993]. *Consciousness Explained*: Penguin Press.

Dennett, D. C. [2001]. Are we explaining consciousness yet? *Cognition, 79*(1-2), 221-237.

Densmore, S., and Dennett, D. [1999]. The virtues of virtual machines. [Article]. *Philosophy and Phenomenological Research, 59*(3), 747-761.

Desimone, R., and Duncan, J. [1995]. Neural mechanisms of selective visual attention. *Annual review of neuroscience, 18*, 193-222.

Edelman, G. [1992]. *Bright air, bright fire: on the matter of the mind*. New York: Penguin.

Fodor, J. A. [1992]. The big idea: can there be science of the mind. *Times Literary Supplement*(July), 5-7.

Gamez, D. [2008]. Progress in machine consciousness. *Consciousness and Cognition, 17*(3), 887-910.

Greenfield, S. [2000]. *The private life of the brain*. New York: John Wiley & Sons, Inc.

Haikonen, P. O. A. [2000]. *Modular neural system for machine cognition.* Paper presented at the Proceedings of the International Joint Conference on Neural Networks, Como, Italy.

Haikonen, P. O. A. [2003]. *The cognitive approach to conscious machines*. UK: Imprint Academic.

Haikonen, P. O. A. [2007]. Essential issues of conscious machines. *Journal of Consciousness Studies, 14*(7), 72-84.

Harth, E. [1993]. *The creative loop: How the brain makes a mind*. New York: Addison-Wesley.

Hawkins, J., and Blakeslee, S. [2004]. *On intelligence*. New York: Henry Holt & Company, LLC.

Hudlicka, E. [2009]. Challenges in developing computational models of emotion and consciousness. *International Journal of Machine Consciousness, 1*(1), 131-153.

Kaiser, W., and Steiner-Kaiser, J. [1983]. Neuronal correlates of sleep, wakefulness, and arousal in a diurnal insect. *Nature, 301*, 707-709.

Kuipers, B. [2008]. Drinking from the firehose of experience. *Artificial Intelligence in Medicine, 44*(2), 155-170.

Leigh, R. J., and Kennard, C. [2004]. Using saccades as a research tool in the clinical neurosciences. *Brain, 127*(3), 460-477.

Liversedge, S. P., and Findlay, J. M. [2000]. Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*(1), 6-14.

Lou, H. C. [1982]. *Developmental neurology*. New York: Raven Press.

Lynch, G., and Granger, R. [2008]. *Big Brain: the origins and future of human intelligence*. New York: Palgrave Macmillan.

Minsky, M. [2006]. *The emotion machine*. New York: Simon & Schuster Paperbacks.

Morsella, E. [2005]. The function of phenomenal states: Supramodular interaction theory. *Psychological Review, 112*(4), 1000-1021.

Nisargadatta. [1973]. *I am that*. Bombay: Chetana Publishing.

Posner, M. I. [1994]. Attention - the mechanisms of consciousness. [Review]. *Proceedings of the National Academy of Sciences of the United States of America, 91*(16), 7398-7403.

Rolls, E. T. [2007]. A computational neuroscience approach to consciousness. [Review]. *Neural Networks, 20*(9), 962-982.

Rosenthal, D. M. [1991]. *The nature of Mind*: Oxford University Press.

Sartre, J. P. [1943]. *Being and Nothingness*. England: Metheun & Co. Ltd.

The Schwarz Foundation: Can a machine be conscious? [2001]. available at: http://www.theswartzfoundation.org/banbury_e.asp

Sloman, A. [1991]. Developing concept of consciousness. [Article]. *Behavioral and Brain Sciences, 14*(4), 694-695.

Sloman, A., and Chrisley, R. [2003]. Virtual machines and consciousness. [Article]. *Journal of Consciousness Studies, 10*(4-5), 133-172.

Spreen, O., Risser, A. T., and Edgell, D. [1995]. *Developmental psychology*. New York: Oxford University Press.

Starzyk, J. A. [2008]. Motivation in Embodied Intelligence *Frontiers in Robotics, Automation and Control* (pp. 83-110): I-Tech Education and Publishing.

Starzyk, J. A. [2010]. Motivated Learning for Computational Intelligence. In B. Igelnik (Ed.), *Computational Modeling and Simulation of Intellect: Current State and Future Perspectives*. Hershey, PA: IGI Publishing.

Sun, R. [1997]. Learning, action and consciousness: A hybrid approach toward modelling consciousness. [Article]. *Neural Networks, 10*(7), 1317-1331.

Sun, R. [2007]. The importance of cognitive architectures: an analysis based on CLARION. [Article]. *Journal of Experimental & Theoretical Artificial Intelligence, 19*(2), 159-193.

Taylor, J. G. [1997]. Neural networks for consciousness. [Article]. *Neural Networks, 10*(7), 1207-1225.

Taylor, J. G. [2007]. CODAM: A neural network model of consciousness. *Neural Networks, 20*(9), 983-992.

Taylor, J. G. [2009]. Beyond consciousness? *International Journal of Machine Consciousness, 1*(1), 11-21.

Velmans, M. [2002]. Making sense of causal interactions between consciousness and brain. *Journal of Consciousness Studies, 9*(11), 69-95.

Velmans, M. [2009]. How to define consciousness: And how not to define consciousness. *Journal of Consciousness Studies, 16*(5), 139-156.

Young, J. Z. [1978]. *Programs of the brain*. Oxford: Oxford University Press.