The Role of Attention in Robot Self-Awareness

Rony Novianto and Mary-Anne Williams

Abstract—**A robot may not be truly self-aware even though it can have some characteristics of self-awareness, such as having emotional states or the ability to recognize itself in the mirror. We define self-awareness in robots to be characterized by the capacity to direct attention toward their own mental state. This paper explores robot self-awareness and the role that attention plays in the achievement self-awareness. We propose a new attention based approach to self-awareness called ASMO and conduct a comparative analysis of approaches that highlights the innovation and benefits of ASMO. We then describe how our attention based self-awareness can be designed and used to develop self-awareness in state-of-the-art humanoidal robots.**

I. INTRODUCTION

Current AI systems, including robotic systems, are unacceptably brittle in the face of change. System designers and developers must explicitly specify all actions and behaviors in order for the system to work as intended such that it achieves its design goals. In open complex and dynamic environments it becomes difficult for systems developers to specify specific actions and behaviors for all possible conditions and situations. Software development approaches to structured and well defined problems do not typically scale in complex and dynamic environments because there is a huge, possibly infinite, number of possibilities that need to be catered for; it is unreasonable to expect that system developers can foresee and develop appropriate responses for all relevant eventualities.

Once deployed, systems are effectively limited by a static set of instructions that encode their designers understanding, conception and perception of the domain in the form of action and behaviors. It is the system designers that conceptualize and model responses to the situations that a robot system will face and kinds of tasks it will be required to achieve. Thus, systems typically fail in circumstances where the developer's design is limited, inflexible, incomplete or incorrect. In addition, system development is a contained process undertaken with limited resources, e.g. funding, time and effort, and as a result the systems developed tend to be limited in scope and capability.

As a consequence, current systems are limited to domain

Rony Novianto is with the Innovation and Enterprise Research Laboratory, University of Technology Sydney, Broadway, NSW 2007 Australia (phone: +61-2-95141940; fax: +61-2-95144535; e-mail: rony@ronynovianto.com).

Mary-Anne Williams is with the Innovation and Enterprise Research Laboratory, University of Technology Sydney, Broadway, NSW 2007 Australia (phone: +61-2-95144503; fax: +61-2-95144535; e-mail: Mary-Anne@it.uts.edu.au).

specific applications where they can perform under a small and finite set of conditions that have been specifically anticipated and represented in a given application. It is not surprising that they fail to perform in open complex and dynamic environments like a typical human world where they can be subjected to unanticipated, unexpected, and unrepresented conditions that can impact them. For example, a typical robot designed for robot soccer performs poorly when exploring in a room rather than a soccer field, because its domain of application is limited to the robot soccer environment and its design goals are also limited to that domain. Even worse, small changes in lighting or playing surface can "disorient" a robot soccer player and reduce their performance.

A self-aware system has the possibility of dealing with novel situations more effectively than a system without self-awareness, because it would have the capacity for introspection that would allow it to inspect and exploit its representations, e.g. internal state.

A self-aware system can attend to its own internal states, thus providing a means of generating introspection and self-modification capabilities. McCarthy describes the importance of mental state introspection for a robot to operate in the common sense world and what kinds of mental states are needed [14]. The common sense world (i.e. human world) contains complex information where self-awareness in systems can enable improved performance and more operational flexibility. Robotic systems have been developed to be aware of their own motion [15], able to imitate [19], [20], driven by emotion [11], [12], and able to change their own models of their physical embodiment [3]. Important work in cognitive robotics in reasoning about action and reasoning about knowledge [18], [13], [7] is relevant. However, there is little work on developing self-aware robots that can *focus attention* on their own internal states in a holistic fashion, i.e. perception and emotions as well as beliefs, which we argue is necessary for genuine self-awareness.

In this paper, we provide a new attentive self-modifying framework (ASMO) for developing an attentive robot with self-awareness based on an architecture that supports the ability of a system to *focus attention on the representation of internal states*.

II. CURRENT STATE OF THE ART

There are some frameworks designed for machine consciousness that also aim to create self-awareness in a machine. However, few are actually implemented in a robot and hence do not provide a convincing demonstration of

self-awareness in a real system. In this section, we discuss the current state-of-the-art in cognitive architectures that support self-awareness and that have been implemented and evaluated in a real physical robot.

Michel, Gold and Scassellati developed an infant-like humanoid robot called NICO that can recognize its own motion in its visual field, including in a mirror [15] (p. 2763). NICO expects to see motion in its visual field whenever certain motor movements commence after a certain time. It learns this time characteristic through experimentation. It labels motions that appear in the visual field within this learned time frame as its own motion, thus it can distinguish itself from others based on the idea of linking motion to time.

The Suzuki, Inaba and Takeno approach to consciousness is to maintain consistency of cognition and behavior of self and others in order to understand the behavior of self and others [19] (p. 101). When a system reaches a state where this behavior of self and others is understood, the system is deemed to be conscious. They determine that the imitative behavior is adequate for analyzing consistency in cognition and behavior of self and others. They conducted four experiments using a robot's imitating actions, including its own actions in a mirror. The result is that the robot passed a mirror test with 70% accuracy [20] (p. 498). Being conscious in this sense, the robot was able to discriminate itself and others much of the time, however the relationship between consciousness and awareness (self-awareness in particular) in this scenario or context is not discussed or clarified. In contrast, we focus on the capability to access and inspect one's own mental state which we refer to as self-awareness (a more detailed definition will be discussed below).

Kawamura *et al.* [11], [12] developed the Intelligent Soft-Arm Control (ISAC) robot that is not self-aware in the sense that it cannot recognize itself in a mirror, but it can deliberate on its emotions based on memory experience. Self-reflection, self-awareness and sense-of-self are represented by a self-agent which consists of a set of agents interacting with memory systems. The emotion that emerges from an activity of experience is learned and stored in memory systems. When an event occurs, emotions activate the episodic memory which in turn activates cognitive control to suppress current behavior and execute required behavior.

Robots have been developed that exhibit an adaptation capability for their own body [3]. These robots can recover from damage or failure that occurs to their body. A robot continuously creates a concept of its own physical structure (self-modeling) and uses this self-model to generate forward locomotion with four legs initially without knowing what its body actually looks like. When the robot's structure changes unexpectedly, it can reform its internal self-model to generate new behaviors to compensate and accommodate these changes. In this case, it remodels the concept of its own physical structure to generate forward locomotion with three legs when one of its legs is removed. This is possible because it has a model of its own physical structure.

A recent robotic system developed for the RoboCup competition plays soccer autonomously, i.e., not remote controlled. The robot soccer players search for the ball; once found, they walk towards it, then grab and kick it towards the opponent's goal [4]. Humans are self-aware when they play soccer and perform these activities; they know what they are doing and why. In contrast, even though the soccer playing robots can perform sophisticated behaviors that would require self-awareness in humans, they can play because they have been designed by humans who have encoded their own knowledge of the game and the required skills in the form of instructions executed as computer programs.

Current approaches to develop self-awareness tend to focus on the performance or abilities of a robot, such as recognizing itself in a mirror, and do not focus on the ability to focus attention on its own internal states. Given most people would accept that mammals are self-aware, and yet not all of them can actually recognize themselves in the mirror, we believe recognition in the mirror is not a valid test, and certainly not a necessary or sufficient condition. Moreover, one can develop a dedicated mirror recognition system without taking self-awareness into consideration. Therefore, we seek to develop a framework that helps explore and develop systems that can focus attention on their representations of internal states. We argue that having an ability to attend to internal states is necessary for self-awareness. Our framework uses this insight to provide a new foundation for self-awareness in real robots.

III. Self-Awareness

Most experts would not consider that a robot is really self-aware just because it can visually recognize its own motion or itself in a mirror since a program specifically designed to achieve that kind of recognition without having a genuine awareness capacity can be developed. The ability to visually recognize one-self is not enough for achieving self-awareness. Self-recognition can be a side-effect of self-awareness, but not a pre-requisite. We believe a robot needs a capability to attend to its internal states in order to be self-aware. Current approaches as described in the previous section do not focus on directing a robot's attention to its own internal processes. If we add an attention process to a robot so that it can focus on processes that happen internally during self-recognition activities then we would consider it to be self-aware. What is crucially important is not the ability to recognize itself in a mirror (e.g. a visual inverted reflection), but rather to be aware of its own emotions, perceptions, beliefs and intentions during the recognition process. If a robot has totally lost all of its outward facing sensations, it may not be aware of its environment (external awareness), however it can still be aware of itself (self-awareness).

At this stage, we focus on a single robot and its ability to

develop an awareness of itself, not others. Clearly we cannot develop an awareness of others in the same way as we can achieve within a single agent like a robot. Just like humans, robots can project awareness of themselves onto others, e.g. we can use feelings to develop empathy for others and can use knowledge about our own consciousness to draw conclusions about the consciousness of others. In other words, we can use our own self-awareness to understand others and to communicate with them.

We define *self-awareness* to be the capability of an agent to focus attention on the representation of internal states. Internal states can be made up of emotion, belief, desire, intention and expectation or it can be processes such as sensation, perception, conception, simulation, action, planning and thought. Humphrey proposes *sentio ergo sum* (I feel, therefore I am) in contrast to Descartes' *cogito ergo sum* (I think, therefore I am), and suggests that sensation is the minimum requirement for being conscious [10].

Perception of internal states may occur without self-awareness. Knowledge can be used to make decisions, to respond or to do further processing with no awareness of the existence of that knowledge. For example, perception of one's belief provides information to make decisions like where to look for an object or to recognize an object, but it is not until attention is drawn to that perception that the robot would be aware he has this belief (self-awareness). Therefore, attention plays a key and defining role in self-awareness. We formalize this concept in our novel ASMO framework, which can be used to design an attention-based self-aware agent.

IV. ATTENTIVE AND SELF-MODIFYING FRAMEWORK

We present a basic framework that can be extended for future development and use (see Figure 1). It includes the following four major components: physical body, perception, self-concept, and attention. This framework is designed so that it can be applied and implemented on
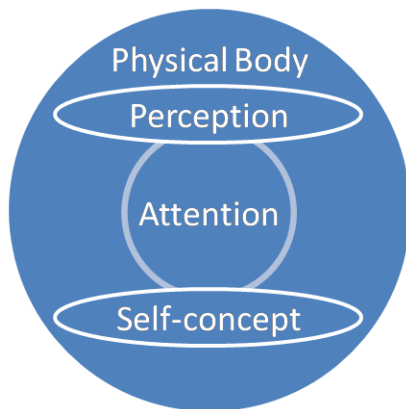

Fig. 1. ASMO framework

different physical robots, i.e. a state-of-the-art Humanoidal Robot called Smokey (see Figure 2) and Nao (see Figure 3). Likewise, human self-awareness can arise in different physical bodies. In this paper, we describe this framework in Smokey.
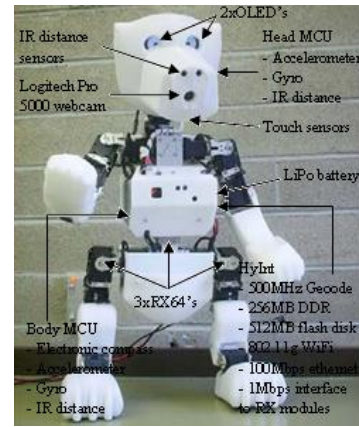

Fig. 2. Smokey

Fig. 3. Nao

A. Physical body

In order to exist in a physical world, a robot has a physical body that can interact with that world. This body consists of sensors that can convert physical stimuli into sensations and actuators that can create physical stimuli in the environment. Physical stimuli exist outside and inside the body whereas sensations always occur inside the body. Physical stimuli are material and have physical characteristics, hence are subject to physical laws whereas sensations are immaterial and do not have clearly physical characteristics even though they exist in the physical body. In humans, although our thoughts are created and reside in our physical brain, they are not considered to be physical objects. Instead, they are ephemeral, immaterial or abstract. Thus, sensation is an immaterial process that executes within a material body. Humphrey's second criteria states that sensations are localized in the body [10] (p. 117).

Outward facing sensation comes from physical stimuli outside the body whereas *inward facing sensation* emanates from physical stimuli inside the body. Sensation related to the visual appearance, the sound, the taste, the smell, and the texture of a body are considered to be outward facing because the physical stimuli captured by the sensors exist outside the body. *Physical sensation* refers to a sensation that converts physical stimuli from either outside or inside the body.

Human exteroceptive senses transform stimuli from outside the body (e.g. our eyes see external objects) while interoceptive senses transform stimuli within the body (e.g. our middle ear for balance). A stimulus outside the body can be a frequency waveform of light reflected from an object in space whereas a stimulus inside the body might be a chemical reaction that triggers a sense of hunger.

B. Perception

Outward and inward facing sensations are processed further to create perceptions. Perception processes (e.g. fusion) the outward facing and inward facing sensation to gain information about self and the surroundings.

Perceiving sensations creates representations about the world and helps determine what is happening in the system

and the world. When attention is focused on the perceptions of the outside world, we say the system has *world awareness of the external surroundings*.

Perception generates internal states and responses can occur, like the evocation of an emotion in response to a perceived object. Although, this emotion belongs to oneself, self-awareness is not necessarily generated immediately. One could remain unaware of one's own emotion, until attention is drawn to it. Thus, self-awareness is generated when attention is focused on internal states. Clearly, attention plays a crucial role in self-awareness.

Perception also processes the inward facing sensations that collect information about the body and create representations of and about the body. It alerts the agent to what is happening to the body. Hence, when attention is focused on this body-oriented perception, awareness about the body (one self's body) is experienced, i.e. *self-awareness* is experienced. Furthermore, inward and outward facing sensations can be integrated to enrich those experiences.

Perception processes sensations to create representations which are interpreted and that can create new knowledge and beliefs through reasoning and learning. This process is an experience and some aspects can be translated into language for sharing. Agents cannot express sensations, but can express certain representations of sensation [24]. For example, it is difficult to express a sensation of red color to someone who is blind since birth, but it is possible to express information about red color such as beliefs, feelings, patterns, and knowledge [24]. Thus, sensation is subjective. According to Humphrey, sensations belong to the subject [10] (p. 115). He suggests that a sensation of what is happening to me and a sensation of what is happening to another person can never be the same.

C. Self-concept

Self-concept contains collections of the facts, conditions, or other representations attained from perception that characterize the self. This characterization gives rise to personality which distinguishes one individual from another individual (individuality). It gives a general representation of "me" which may include what I "think" other individuals "think" of me. Perception affects the representations of the surroundings (world-conception) and self (self-conception) while perception itself is also affected by them.

The representations derived from perception can be "consistent" or "inconsistent" with self-concept. Self-concept grows clearer and stronger when it is consistent. One has more confidence of what one thinks about oneself when one's perception is consistent. When it is inconsistent, perception can modify the self-concept: changes may occur instantaneously or over periods of time. For example, when one attends to an internal belief and discovers that it is inconsistent with old beliefs, the beliefs may change instantly or may require a lifetime of consideration depending on the nature of the inconsistency [2].

Self-awareness as defined earlier is needed to foster and direct these changes. ASMO allows a robot to be attentive to these changes and also the representation of its self-concept.

D. Attention

Many have suggested the importance of attention in self-awareness [8], [16], [6]. Attention can highlight representations, e.g. beliefs and processes, for an agent in being unaware to aware. For example, before reading this sentence, you were probably not aware of the sensations in your feet although these sensations were happening. When attention is focused on them, these sensations are perceived and awareness is shifted to the sensations. Attention guides awareness.

Attention can be directed voluntarily and involuntarily. One can direct attention to internal states voluntarily without following designated sequences or procedures, e.g. one can change attention from belief to emotion to intention in any order. Involuntarily attention happens during an occurrence of an event or a physical stimulus either inside or outside the body. In this case, attention is often directed to the most salient states. For instance, someone asking how we feel can trigger attention to be directed towards our feelings. If a stimulus that triggers hunger occurs at the same time, attention may choose between emotional or hunger state according to the most salient.

Voluntarily attention is an action of one's will – where an agent can exercise choice and control. Other individuals can influence one's attention, but they cannot control it. Voluntarily attention can influence involuntarily attention, and vice versa.

Attention for self-awareness is like a program counter for a central processing unit (CPU) that indicates where the process is happening right now. Thus, attention provides a sense of awareness by providing knowledge about what is happening right now. If it is directed to the internal states, then it gives a sense of self-awareness. It is directed at one state at a time – while humans can perceive multiple things at once, humans cannot think about two things in any given instant [1] (p. 99), e.g. consider adding and multiplying two numbers simultaneously.

An agent can achieve virtual awareness of more than one thing or task at a time because attention can switch back and forward between several tasks rapidly, just like multi-tasking in contemporary operating systems. When one task demands more attention than others, the performance of the other tasks may be affected. An agent appears to be able to do several tasks without them affecting each other with practice, even though it is attending to one at a time.

V. EVALUATION

We have been implementing our approach and framework on the state-of-the-art prototype humanoidal robot Smokey which has a humanoid form consisting of a head, neck, torso, two arms (elbow-like joints) and two legs (knee-like joints) [22], [5]. Currently it has monocular vision based on only one camera for vision. However, due to its extensible design and standard component engineering, an additional

camera could be added in order to achieve binocular vision with two cameras in the future. Our aim with Smokey is to empower him as an attentive self-aware robot that can plan, communicate and collaborate in open complex and dynamic environments. The following scenario has been designed to help test and demonstrate this attentive capability and self-awareness based on definition given earlier. Smokey lives in a real world where he interacts with people and real objects. He plays the bongo drums, which are his favorite musical instrument. If he is happily playing the bongos when a person comes along who takes the drums away, then he feels "unhappy". He "cries" and asks the person to return the drums to him. Smokey is simulated by the act of a "*removing the drums*" and in response, he suppresses his current "*bongo playing behavior*" and activates the "*crying/requesting behavior*". He is able to select activities that are more significant and meaningful to him as determined by his sense of self. This scenario is similar to the ISAC scenario where ISAC asks people to leave the room when a "*fire*" stimulus is detected [11], [12].

However, we argue that Smokey does not demonstrate genuine self-awareness above because emotion can arise in an agent and it can behave consistently according to the emotion but not be self-aware unless attention is actually focused on the emotion. In the scenario above, Smokey's "*crying/requesting behavior*" is *automatically* driven or triggered by his "feeling" based on an encoded emotional response. In other words, Smokey's behavior is an automatic triggered response.

An *automatic process* has three key characteristics: it occurs without conscious awareness, without intention, and requires little or no attention [17]. Therefore, Smokey's automatic behavior as described above does not demonstrate self-awareness. If he can focus his attention on his "unhappy feeling", then he would satisfy our definition of self-awareness.

A self-aware system can execute processes that are not designed or preprogrammed for the current ambient stimuli. Without self-awareness, Smokey would continue to perform the "*crying/requesting behavior*" whenever the same stimulus occurs, i.e. the drums are taken away. Young children often respond and perform similar behavior when someone takes something from them. Armed with self-awareness though, an agent can respond differently to the same stimulus. Given that Smokey's intention is to be a "happy" robot, he can suppress the "*crying/requesting behavior*". He can activate "*informed behavior*" which indicates that he has finish playing the bongo drums, and although he may be "unhappy" because someone took his bongos away, he may chose to do something else. This behavior reflects a similar increase in the level of maturity where an older child can ignore automatic responses better than a younger child. Human self-awareness is not static. It develops and matures over time. Self-aware systems can suppress automatic or innate processes that are driven by existing stimuli: they can ignore or select controlled processes based on consistency with their intentions and self-concept.

Our experimental scenarios are based on the notion of attention which is used to suppress, override and select responses. The scenario we use to evaluate attention based self-awareness can be described as follows:

1. Smokey plays the bongo drums whenever they are offered
2. A person comes by and takes away the drums
3. Smokey becomes unhappy, cries, and requests the drums back
4. Smokey becomes aware that his tantrum is not achieving the desired goal of having the drums returned.
5. Smokey can stop crying and requesting the drums back.
6. Smokey can inform the person that he has finished playing the drums and do something else.

This attention based self-aware system can be implemented in a way that allows Smokey to modify his representation or even his own code. When the "unhappy feeling" is evoked by the situation where the drums are taken away, attention may or may not be directed at this process straightaway depending on the competing stimuli and states. Once it is directed to this process, Smokey can perceive his feeling which leads to awareness of it. This mechanism provides higher order perception to deliberate and re-plan his behavior according to his intention and self-concept. This attention indicates what and where his representations or the code needs to be modified. The actual code modification could be carried out by human designers or Smokey himself, depending on the complexity of the required change. The key idea is that Smokey himself identifies the processes or the behavior that need to be changed or evoked differently. Self-modifying robots are more adaptable to open complex and dynamic environment, because they can change response on the fly.

In classical robotics approaches, Smokey's behaviors and representations are based on the developers' understanding of the design goals and the domain. A system that can modify itself has the chance to change according to its own desires, intentions and beliefs and hence one could argue that it is more grounded [21]. Grounding is crucial for performance and intelligence in artificial systems [24]. Developers and other people can influence Smokey to change his behavior by providing him with different stimuli, but Smokey still makes the decisions himself depending on whether the stimuli are strong enough to make him want to change or not. Two identical self-aware robots can behave differently depending on the way they interact with the environment. This view is shared with developmental robotics community [23].

VI. COMPARATIVE ANALYSIS

We compare our ASMO framework with five other relevant frameworks: the motion-based self-recognition

system (NICO) [15], the mirror image cognition system (MIX) [19], [20], the self-agent (SA) [11], [12], the self-modeling system (SM) [3] and WrightEagleUnleashed RoboCup system (WEU) [4]. We limit our comparison to these frameworks, because they have been implemented on a real robot. We use binary types of symbols: "O" to indicate that a criterion is specifically supported and "X" to indicate otherwise. We also use categorical types of symbols "[ ]" for "not explicitly supported", "[-]" for "partially supported" and "[+]" for "strongly supported".

A. Recognition

1) Motion recognition

This criterion involves the ability to recognize one's own motion. NICO recognizes a wide range of movements of its own motion whereas MIX recognizes only forward, stop, and backward movements. Other frameworks do not recognize their own motion. Therefore, NICO gains "strongly supported", MIX gains "partially supported" and the rest gain "not explicitly supported".

2) Mirror recognition

NICO can recognize its own motion in the mirror in addition to the direct visual recognition. MIX differentiates between its own image in a mirror and other robot's images who imitate it. Other robotic systems in our comparison cannot recognize themselves in the mirror. Thus, "O" for "supported" is given to NICO and MIX and "X" for "not supported" is given to others.

3) Imitation recognition

This is the ability to distinguish self from imitations. MIX can discriminate between the visual image of itself and other robots either when the other robots "intentionally" imitate it or when other robots directly copy its movements via cable. Other frameworks do not directly support this kind of imitation recognition performed by other robots. Hence, MIX receives "supported" and others receive "not supported".

B. Model

1) Emotion model

SA uses an emotion agent to model emotion that can activate the episodic memory from a stimulus in the environment. ASMO currently does not have an emotion model. Because emotion is one of its internal states, it has the structure to notice emotion via attention when this kind of complex emotion model is available. We consider this kind of emotion model to be the next step in the development of ASMO, which we expect to develop in the near future, because one case of the experimental scenario for Smokey requires dealing with emotion. Other frameworks do not have an emotion model. Thus, "O" for "supported" is only given to SA.

2) Memory model

SA integrates three different memory systems that correspond to humans: short-term memory, long-term memory and working memory. The long term memory consists of procedural, episodic, and declarative memory.

NICO uses a memory module to implement simple object permanence. It remembers the object in its visual perceptions and assumes that the object still exists when it disappears from its vision for couple of frames.

WEU uses many kinds of memory to perform complex and sophisticated behaviors too. However, it does not explicitly identify short-term, long-term, and working memory.

Likewise, ASMO does not explicitly distinguish between short-term, long-term, and working memory. MIX and SM do not integrate memory into their systems. Hence, SA gets "strongly supported". NICO, WEU, and ASMO get "partially supported". MIX and SM get "not explicitly supported".

3) Physical body model

SM develops its own physical body model that can be used to generate behaviors when some changes happen to its body. Other systems do not have self-modeling of physical body model implemented in their robots. Therefore, only SM obtains "O" for "supported" and others obtain "X" for "not specifically supported".

C. Process

1) Sophisticated behavior

WEU performs sophisticated but completely specified and automatic behaviors used in the RoboCup Robot Soccer competition. Other systems perform simple arm motion movements, locomotion, and experimentations. Therefore, WEU achieves the only supported system for this criterion.

2) Attention to internal state

SA does not describe the use of attention [11]. The edited version of the same paper available online [12] includes the use of attention in working memory system and central executive agent to select information [12] (p. 4) and an

TABLE I
UNITS FOR MAGNETIC PROPERTIES

Criteria	NICO	MIX	SA	SM	WEU	ASMO
Motion recognition	[+]	[-]	[ ]	[ ]	[ ]	[ ]
Mirror recognition	O	O	X	X	X	X
Imitation recognition	X	O	X	X	X	X
Emotion model	X	X	O	X	X	X*
Memory model	[-]	[ ]	[+]	[ ]	[-]	[-]
Physical body model	X	X	X	O	X	X
Sophisticated behavior	X	X	X	X	O	X
Attention to internal state	X	X	X^	X	X	O
Self-modifying code	X	X	X	X	X	O

action [12] (p. 5) respectively. However, it does not describe how this kind of attention can be used to focus on perception of internal states. ASMO is designed to achieve this criterion so it supports attention to its internal states. Thus, ASMO attains the only "O" while the other systems receive an "X".

3) Self-modifying code

ASMO has an attention system that indicates what representations are being enacted and what code of a process is currently executing. It can be aware of the internal states, so it has the potential to self modify this code to suit more complex conditions and adapt to its dynamic environment. No other framework has this ability to change its own codes. Although the robot used in SM can compensate for changes that happen to its body, it uses 15 predefined candidates of body model to match to the best of its environment [3]. There is no facility to identify or change the relevant code. Hence, only ASMO attains "O" for "supported".

VII. Conclusion

Self-awareness is a high order capability and an important requirement for humanoids. If humanoids are to communicate and collaborate effectively with humans, then they need to be self-aware and to modify their representations all by themselves in order to exhibit genuine intelligence [24]. In the common sense human world, robots will need to be able to focus their attention on what is happening to and within them. We argue that attention directed towards internal states is necessary for self-awareness. We introduce a new framework ASMO that supports self-awareness as the ability to focus attention on the perception of internal representations like states. A robot based on ASMO can identify and self-modify its own representations via introspection of its own internal states and deliberation which can lead to new more flexible behaviours. We are implementing this framework on our state-of-the-art humanoidal robot Smokey and later a Nao in such a way that it will provide them with an attentive capability that can be used to gain awareness of relevant aspects of himself and his environment.

Acknowledgment

References

[1] J. R. Anderson, *Cognitive psychology and its implications*. Ed. 6, New York: Worth Publishers, 2005.

[2] S, Benferhat, S. Kaci, D. L. Berre, and M-A. Williams, "Weakening conflicting information for iterated revision and knowledge integration," *Artificial Intelligence*, vol. 153, no. 1-2, pp. 339-371, Mar. 2004.

[3] J. Bongard, V. Zykov, and H. Lipson, "Resilient machines through continuous self-modeling," *Science*, vol. 314, no. 5802, pp. 1118-1121, Nov. 2006.

[4] X. P. Chen, H. He, B. Johnston, F. Liu, A. Muh, R. Novianto, C. Stanton, F. Wang, X. Wang, M-A. Williams, K. Xu, and L. Yuan, (2008, August 10). RoboCup 4-legged SPL system. Available: http://www.wrighteagleunleashed.org

[5] M. Dickinson, and P. Turner, "Development of a biomimetic robotic bear: or is a bare bear bearable?," in *IEEE Int. Conf. Robotics and Biomimetics*, Bangkok, 2008, pp. 7-12.

[6] T. S. Duval, and R. A. Wicklund, *A theory of objective self-awareness*. New York: Academic, 1972.

[7] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi, *Reasoning about knowledge*. MIT Press, 2003.

[8] P. Gardenfors, *How homo became sapiens*. Oxford: Oxford University Press, 2003.

[9] P. Gardenfors, and M-A. Williams, "Reasoning about categories in conceptual spaces," in *Proc. 17th Int. Joint Conf. Artificial Intelligence*, Seattle, 2001, pp. 385-392.

[10] N. Humphrey, *A history of the mind*. London: Chatto & Windus, 1992.

[11] K. Kawamura, W. Dodd, P. Ratnaswasd, and R. A. Gutierrez, "Development of a robot with a sense of self," in *Proc. 6th IEEE Int. Sym. Computational Intelligence in Robotics and Automation*, Espoo, 2005, pp. 211-217.

[12] K. Kawamura, W. Dodd, P. Ratnaswasd, and R. A. Gutierrez, "Development of a robot with a sense of self," unpublished. Available: http://eecs.vuse.vanderbilt.edu/CIS/pubs/2005-Kawamura.et.al-CIRA.pdf

[13] H. J. Levesque, R. Reiter, Y. Lespérance, F. Lin, and R. B. Scherl, "GOLOG: A logic programming language for dynamic domains," *The Journal of Logic Programming*, vol. 31, no. 1-3, pp. 59-83, Apr-Jun. 1997.

[14] J. McCarthy, "Making robots conscious of their mental states," in *Machine Intelligence*, vol. 15, S. Muggleton, Ed. Oxford: Oxford University Press, 1996, pp. 3-17.

[15] P. Michel, K. Gold, and B. Scassellati, "Motion-based robotic self-recognition," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Sendai, 2004, pp. 2763-2768.

[16] A. Morin, "Levels of consciousness and self-awareness: A comparison and integration of various neurocognitive views," *Consciousness and Cognition*, vol. 15, no. 2, pp. 358-371, June 2006.

[17] M. I. Posner, and C. R. R. Snyder, "Attention and cognitive control," in *Information Processing and Cognition*, R.L. Solso, Ed. NJ: Erlbaum, 1975, pp. 55-85.

[18] R. Reiter, *Knowledge in action: logical foundations for specifying and implementing dynamical systems*. MIT Press, 2001.

[19] T. Suzuki, K. Inaba, and J. Takeno, "Conscious robot that distinguishes between self and others and implements imitation behavior," in *Innovations in Applied Artificial Intelligence*, vol. 3533, Heidelberg: Springer, 2005, pp. 101-110.

[20] J. Takeno, K. Inaba, and T. Suzuki, "Experiments and examination of mirror image cognition using a small robot," in *Proc. 6th IEEE Int. Sym. Computational Intelligence in Robotics and Automation*, Espoo, 2005, pp. 493-498.

[21] M. Trieu, and M-A. Williams, "Grounded representation driven robot motion design," in *RoboCup 2007: Robot Soccer World Cup XI*, vol. 5001, Heidelberg: Springer, 2008, pp. 520-527.

[22] P. Turner, and M. Dickinson, "Hykim – development of a robot bear: bringing the strength and robustness of a bear's biomimetic features to a robot," in *IEEE Int. Conf. Robotics and Biomimetics*, Bangkok, 2008, pp. 13-18.

[23] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, "Autonomous mental development by robots and animals," *Science*, vol. 291, no. 5504, pp. 599-600, Jan. 2001.

[24] M-A. Williams, "Representation = grounded information," in *Proc. 10th Pacific Rim Int. Conf. Artificial Intelligence*, Hanoi, 2008, pp. 473-484.