Proceedings of International Joint Conference on Neural Networks, San Jose, California, USA, July 31 – August 5, 2011

# Mental Saccades in Control of Cognitive Process

Janusz A. Starzyk, Senior Member, IEEE

*Abstract*—**This paper proposes a cognitive architecture that uses mental saccades to perform cognitive search in support of motivated behavior and learning. It is intended to control the behavior of robots in real environments and avatars that learn how to operate in virtual worlds. A mental saccade is a parallel concept to the visual saccade and yields a sequential cognitive search for the most likely solution to a problem. This model uses an attention switching mechanism that combines the effect of observations, internal motivations and abstract cognitive planning. Thus, a system that uses this model, will not only follow its internal motivations but will also take advantage of opportunities that present themselves in the environment. This model is intended for development of computational cognition, learning and intelligence in a machine.**

## I. INTRODUCTION

PRESENTED in this paper is a structural approach to control of the cognitive process by using a new concept of mental saccades. Mental saccades are used to perform sequential cognitive searches of the observed scene and associative memory, while responding to the machine's needs and goals. These sequential searches are compatible with cognitive architectures that use a symbolic approach to represent concepts, except in this work symbols are replaced by distributed representations obtained in a network of neurons responsible for perception. These distributed representations of the perceived objects and actions are grounded and the same group of neurons are activated once the same object is recognized or the same action is initiated.

The main aim of this model is to equip the machine with a mechanism that allows it to perceive, plan, and execute motor control in a natural environment. This model uses grounded representation of objects, motivations, and motor control based on distributed activation of neurons in associative memory models. It employs the neural blackboard architectures discussed by van der Velde and de Kamps [1] to obtain these grounded distributed representations. Blackboard architecture solves fundamental problems of cognitive neuroscience of massive bindings, symbol grounding with multiple instances of the same object category in complex scenes, variable bindings in working memory with binding in long term memory. As in the blackboard approach, representations are situated in specific areas of associative memory and manipulated through combinatorial structures to obtain the desired mental process and related action control. To provide grounding, the neural structures that encode the concept are embedded in the associative memories and their duplication to other memory areas is difficult. Thus, manipulations of these concepts are organized by dynamical combinatorial connection structures as well as gating and memory circuits that follow the overall blackboard strategy.

However, the approach presented in this paper goes beyond the combinatorial structures proposed in the blackboard architectures, and while being compatible with the blackboard approach, extends its utility towards systems with motivations, attention switching and action planning.

## II. RELATED WORK

Cognitive architectures are defined as computational processes that are capable of cognition, learning, reasoning, and intelligent control. The term architecture relates to either production rules or the interconnection structure that processes incoming information to control a machine's behavior. They are either based on symbol manipulation, emergent interactions, or a hybrid combination.

Well known examples of symbol manipulation architectures use various design principles like rule-based architectures (SOAR [2]), declarative and procedural knowledge based on if-then rules (ACT-R [3]), logical networks (SNePS [4]), or goal oriented reactive skills (ICARUS [5]). Although these systems can learn new rules, organize their memory or interface with perceptual modules for interface with the real world, their symbol manipulation is not compatible with processing by the brain, where symbols are not precisely defined, interpretation of their properties and associations can change, and memories fade gracefully.

Another group are emergent cognitive architectures that operate on the principle of neural network processing, where symbols emerge from low-level interactions between neurons. Examples of these architectures include NuPic based on hierarchical temporal memory [6] for real-time perception, IBCA [7] that models brain architecture using hierarchical processing and distributed representations of concepts, Cortronics based on confabulation theory [8] that learns to anticipate the next input, and NOMAD [9] based on neural Darwinism used for real-time pattern recognition. So far IBCA has been limited to small scale networks without a clear path towards higher level cognition and action control.

The third group of cognitive architectures contains hybrid architectures that combine neural based low level perception with symbolic processing used for performing high-level cognitive functions. Examples include: CLARION [10] that provides an interface between symbolic and sub-symbolic

Manuscript received January 31, 2011.

J. A. Starzyk is with the School of Electrical Engineering and Computer Science at Ohio University, Athens, OH, USA (phone: 740-593-1580; fax: 740-593-0007; e-mail: starzykj@gmail.com).

978-1-4244-9637-2/11/$26.00 ©2011 IEEE 

processes, DUAL [11] that supports an emergent architecture with interacting micro-agents, Polyscheme [12] where various "specialists" cooperate to guide the machine, and Open Cog [13] designed for virtual agent control. The major problems for hybrid architectures are building interfaces to symbols from low level representations as well as processing uncertain information. An excellent overview of cognitive architectures is presented in [14].

Three architectures are particularly relevant to the approach presented in this paper: LIDA [15] based on global workspace theory with cognitive cycles, a neural blackboard architecture [1] that grounds cognitive cycles on distributed neural representations, and MicroPsi [16] that uses urges for motivations to control machine's actions and learning.

In LIDA, the system operates by constantly extracting and processing cognitive cycles. Such cycles combine elements of attention, perception and action. An example of a cognitive cycle is that of an animal sensing the world, comparing sensation to memory, and choosing an action. Cognitive assemblies compete with each other and the winner is sent to procedural memory for possible action. In LIDA the agent broadcasts information about its focus of cognitive attention to other parts of the system memory (including episodic and semantic memory) that interact in the global workspace. Procedural memory uses templates that best match the cognitive assembly for possible actions. It attempts to estimate the likelihood that the action suggested by the winning assembly will succeed. Copies of the matching templates are made and instantiation of a template is passed to the action selection mechanism. The selected action is executed and the resulting changes in the environment are observed.

MicroPsi, based on the theory developed by Dietriech Dorner [16], uses demands and cognitive urges to control a machine's behavior. Developed by the author of this paper, the theory of motivated learning (ML) uses pains to motivate a machine's actions. While primitive pains are predefined and correspond to demands in [16], a motivated machine is capable of creating abstract pains that are similar but differ from cognitive urges. Like MicroPsi, ML reduces uncertainty about the environment. The major difference is that abstract pains are the result of interaction with the environment, thus any needs that a machine develops are naturally developed, unlike less specified cognitive demands in [16] like competence and social affiliation. Rather than maximizing the external reward, the MicroPsi architecture minimizes its urges. The same overall objective drives ML.

This paper introduces a natural mechanism for cognitive observation, attention switching, action selection and control based on the ideas of ML and global workspace theory.

## III. VISUAL AND MENTAL SACCADES

This section presents the mechanism of mental saccades that uses the idea of a global workspace [17], [18]. The idea of mental saccades parallels that of visual saccades in visual perception. Visual saccades result in scanning a visual scene providing an inspection focus. Visual saccades are also used to study motor control, cognition and memory in functional imaging and transcranial magnetic stimulation [19]. In [20], it was observed that saccadic eye movement data are a reflection of the cognitive process during visual search and reading and can be used to measure its efficiency. It was determined that "where" and "when" aspects of saccades are separate and depend on attentional selectivity [21] and semantic properties of the fixated objects [22].

In this work a similar approach is used to scan the global workspace [17] for memories associated with the perceived objects. This scanning of the global workspace is defined as mental saccades. In the proposed model, mental saccades constitute a core mechanism for conscious selection of the attention spotlight, providing an engine for the machine's conscious behavior.

To illustrate mental saccades, let us consider Fig. 1. It shows the activation of memory traces in the frontal cortex area (global workspace) based on the observed scene. An input image is shown in Fig. 1 together with a selected part of this image (a male figure) that corresponds to the visual focus area resulting from a visual saccade. Upon recognition of the object in the visual focus area a grounded representation of this object becomes an attention spotlight. In this case, a corresponding area of the face recognition is highly activated by feedback from the frontal cortex and associated areas of semantic and episodic memory that relate to the recognized person are primed for activation. These in turn activate memory traces in the global workspace area that will be used for mental searches (mental saccades). Symbolically, the recognized person and primed areas of the working memory are represented in Fig. 1 by memory traces in the frontal cortex.

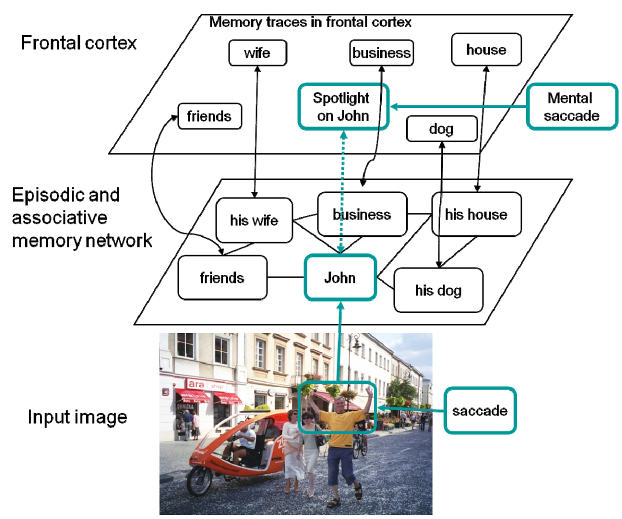


Fig. 1. Selection of the cognitive spotlight

Activated memory traces in the global workspace area are searched in a process similar to visual saccades. For instance, in the illustrated activation in Fig. 1, a mental saccade may proceed from the observed person (John) to one of the concepts associated with this person, e.g. his dog. Thus the attention 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 that was put in attention spotlight 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 primed areas of the working memory are searched through.

Mental saccades constitute an alternative solution to the idea of attention codelets competing for consciousness as presented in the LIDA model [23]. What differentiates this approach from the LIDA model is that it uses an explicit mechanism that switches machine attention to grounded representations that can activate memory traces in the global workspace. In particular, this mechanism does not refer to ill-defined cognitive "atoms" or "attention codelets" needed to begin the consciousness phase in LIDA. Also avoided are references to vague terms like "contents of consciousness" used to describe the conscious state of mind. In this model there is no movement of selected portions of the current situational model to the global workspace required in the LIDA model. A situational model (as discussed in LIDA) or its portion would be difficult to isolate from complex interactions between motivations, observations and inner thoughts. In addition, saccadic search is preferred here over direct competition between primed memory areas, since it is not clear how unconscious perceptions may compete with each other in the global workspace. However, the two theories are close since both are bringing a selected cognitive experience into the attention spotlight for further processing.

Mental saccades are not only evaluating internal thoughts against needs and motivations but also against internally posed questions like "what does a giraffe have for dinner"? These internally posed questions complement motivation driven behavior. For example, you are waiting in the doctor's office (so your goal is to see the doctor), yet since you have nothing else to do you ponder. Your internal talk is not driven by your needs at the moment. For instance, you were thinking about your friend and recalled that his birthday is next week. This may motivate you to remember that you need to send him a birthday card. Although the idea of sending him a birthday card was not driven by your dominant motivation, it did satisfy your need to maintain friendship. In this example, your plan to act was a result of an opportunistic match between a concept of sending him a card and one of the lesser needs to maintain friendship.

Motivations may also drive inner questions. If you feel thirsty you may ask yourself, where can I find a convenience store to buy a drink, and this may drive your visual saccades to find one. For instance, you may ask looking at the picture on Fig. 1 – is there a coffee shop on this street? And this question will drive your visual saccade to help find an answer. The nature of this question may indicate familiarity with the scene and type of environment as well as possibility that the answer is positive. You will most likely not ask this question looking at the desert, but instead you may be interested to see if there is any vegetation or water nearby.

You may also want to remember your thoughts, in case you need to consider another option that you did not chose previously. For instance, if you thought that it will be nice to find a coffee shop, and then you switched your attention to a nice looking women walking by, you would like to be able to come back to your previous thought about the coffee shop. This would require a short term episodic memory that will store your previous thoughts in addition to the working memory that preserves your current thoughts. Thus, your conscious thoughts may work like real perceptions, triggering other thoughts and memories, directing you to a given target, and changing motivations. Most likely an episodic store of your inner thoughts will carry less emotional weight than real episodes, so their storage time will be much shorter than that of real episodes. An exception here may come if you stumble across an important discovery, solve a significant problem, or compose a poem – the thoughts you would like to recall, organize, and use later.

At a lower level of mental development, this inner talk will be based on associations between perceptions and actions, rather than words and meanings. At a higher level, it will use language to formulate problems for investigation and to define objectives for such mental searches.

Of course, an important question is what mechanism is behind your thinking process and what drives questions in the inner talk. Next, the proposed organization of the cognitive process based on mental saccades and attention switching will be described.

## IV. COGNITIVE PROCESS

Fig. 2 illustrates the major working memory organization blocks and flow of information to use mental saccades for attention focus and conscious mental process. Whereas it is not hypothesized 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.

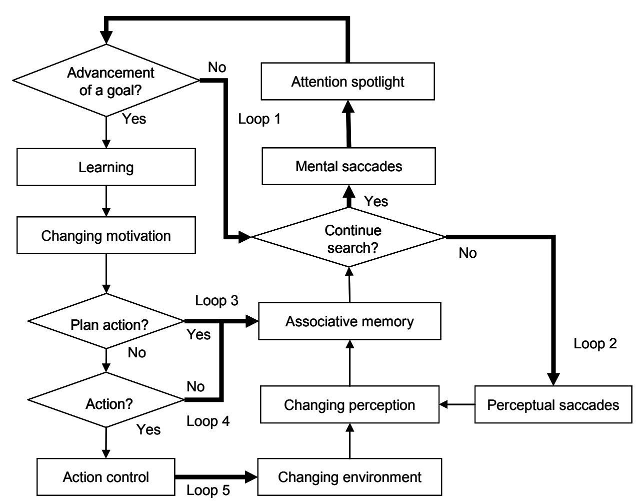


Fig. 2. Mental saccades in a cognitive machine model

Let us trace the major loops in Fig. 2 to explain possible outcomes of a cognitive cycle and creation of dynamic combinatorial memory structures. These memory structures result from changes in the environment, internal motivations and mental probing of alternative solutions to a selected problem that drives the machine's behavior.

Mental saccades constitute the innermost, most frequently updated loop of attention switching (Loop 1 in Fig. 2). They activate a selected memory trace and bring the associated perception to a conscious attention spotlight. The 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 mental saccade cycle, 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. In such case, a new goal can be 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 the internal state of the machine. This innermost "mental" loop compares various alternative solutions corresponding to action planning. It does not have to lead to the implementation of an action by activation of motor control functions. Instead, this innermost loop can lead to a "theoretical solution" that is conceptualized and perhaps stored in the episodic memory for later use.

Loop 2 is associated with changing perceptions due to visual saccades. Such changes may occur either voluntarily, when conscious examination of a given mental scene is completed or as a result of an attention switch (conscious or subconscious). In an advanced cognitive system working in familiar environment, mental saccades may be more frequent than visual saccades. Otherwise, they may happen during a mental saccade, interrupting the thinking and planning process.

If the idea in the attention spotlight was contemplated to advance a current goal, then a decision is made as to 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 the attention spotlight one by one within the third loop (Loop 3). A cycle in Loop 3 is completed by bringing a new scenario into associative memory, which starts a new series of mental saccades related to this new step in a conscious search for a solution.

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 a possible update of sequential memory resulting from learning the results associated with the mental task completed in this loop.

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 visual and mental saccades.

In addition to the typical flow of the conscious process described here, changes in the environment that result from execution of a task or external interference automatically influence the machine's perception and indirectly affect its conscious process. In fact, if it is desired that a machine focuses on a mental task that requires some form of isolation from the 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 a critical element of ML [24] influence the cognitive cycles by changing either their motivations or the content of the working memory. What is important is that between the changes in the environment and changes of the internal state of the machine, one can be almost certain that the cognitive process will never stop.

In what follows higher level architectural structures that can implement the ideas described in the mental saccades model will be presented. They are based on grounded object representations in neural networks that identify what object is observed and where it is located in the visual space [25].

## V. COGNITIVE ARCHITECTURES

### A. Visual Saccades

Feedforward neural networks can perform object identification [26] using distributed representations that represent groups of features (like shape, color, orientation). These networks were modeled after neural activities observed in the ventral stream in monkeys during object attention tests [27]. The ventral stream that contains the anterior inferotemporal (AIT) cortex is involved in object recognition. In a dorsal stream a separate region known as the lateral intraparietal cortex (LIP) was activated depending on the location of the perceived object. The two regions are associated such that the animal could saccade directly to the desired object in experimental tests.

Following this approach van der Velde and de Kamps developed a neural network model for feedforward object identification to obtain grounded object representation and object location [25]. They used top-down feedback links to activate a desired object representation. They demonstrated that object location can be retrieved by association to the ventral stream responsible for object identification. The opposite association is also possible – by priming a specific location, a saccade eye movement to the target object at this location can be generated and the object recognized. A simplified organization of the neural network that provides grounded object representation and its location is shown in Fig. 3.

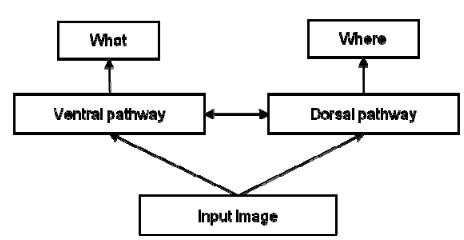


Fig. 3. Visual cognition of what and where

In this section the mechanism described in [25] will be used to show how saccade movements to various target objects can be organized based on object locations activated in the dorsal stream. It is a simple model that does not involve the cognitive aspects of vision when a saccade is directed to the part of scene where an interesting perception is expected.

In this simplified model, visual saccades start with selection of the most salient feature observed in the input image. Activation of the place cells in the dorsal pathway is associated with this particular location. Although we may frequently reexamine the same location in the image, in this model we will assume that after inspecting an interesting place, the saliency signal of the image in this location is lowered. This will allow initially less salient feature to be examined.

Consider the image presented in Fig. 4 with several salient features to be examined (A, through D).

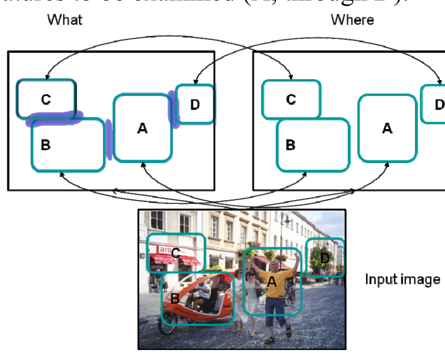


Fig. 4. Visual saccades to salient features.

After inspecting object at location A, the visual perception is focused on location B, C, and D, each time invoking memories and triggering the mental saccades.

Suppose that as a result of a visual saccade to location A in Fig. 4 a grounded representation of my friend (John) was activated in a semantic associative memory, and through associations, other representations related to my friend were primed in the associative memory, activating primed memory neurons E,F,G,H,I (our common friends, his wife, their business, his house and his dog) in the working memory. In the next section how mental saccades are organized in this model will be discussed.

## B. Mental Saccades

A single representation is selected in semantic memory for consideration using the mental saccade mechanism. This mechanism selects primed memory areas based on neural activity level, and applies inhibition to other activated areas for the duration of a mental saccade. A simple mechanism based on winner-take-all (WTA) competition determines which primed area will be considered. The winner of the competition becomes the cognitive attention spotlight, and by using feedback to associative memory the selected representation is highly activated. This corresponds to broadcasting the cognitive spotlight in the global workspace memory discussed by Baars [28] without copying the activated representation of object A to other parts of memory. As in the global workspace proposed by Baars only one representation can be dominant at any given moment, leading to a sequential nature of cognitive processing. Once an attention spotlight is selected by feedback from the working memory, its associated memory areas are also primed for further analysis. This analysis will be performed using mental saccades responsible for attention switching.

Network level organization of the mental saccade mechanism is shown in Fig. 5. Associative memory primes neurons in the working memory of the frontal cortex as illustrated in Fig. 1, and this activation of the working memory neurons is represented by activation of the corresponding primed memory neurons in Fig. 5. These neurons compete using WTA and a single activation is transmitted to dual neurons. Dual neurons remember previous winner activations. A “next mental saccade” signal is used to force the mental saccades. Once the next mental saccade signal is activated, a current winner is inhibited by activating “winner inhibition neurons” (WIN). WIN neurons use the combined inputs from previously activated winners and next mental saccade signal. Inhibition from WIN neurons lowers activation of the primed memory neurons that were previous winners allowing a new winner to be selected. So if the second most active area in the associative memory was related to primed memory neuron H and neuron A was inhibited by its WIN neuron, then H will be selected as a new winner and the attention spotlight will be directed towards associative memory area H. This will be accomplished through feedback from activated primed memory neuron H towards associative memory representation neurons.

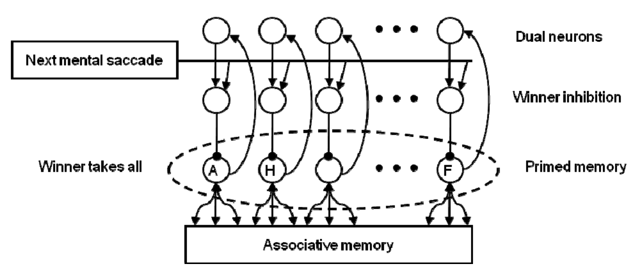


Fig. 5. Mental saccade network model

Notice that once neuron H is selected as a winner of the new competition, the attention spotlight moves towards a concept not observed in the visual input, the object whose representation is recalled from primed associative memory represented by H in the global workspace. Once this new representation becomes the attention spotlight, it may activate other memory areas that are associated with H, such that a mental saccade can move away from the observed object to objects or memories that are related to it or its close associates. Mental saccades continue until all significant memories are revisited or a goal was attained and the attention spotlight moved somewhere else.

Since activation of dual neurons in the mental saccade mechanism will fade away after several seconds we may be able to revisit the same idea again (perhaps with different context). The question is what happens after various objects from memory are brought into the attention spotlight by this mechanism of mental saccades. This is discussed next.

### *C. Goal advancement, learning and planning*

Let us first consider an architecture that will generate the next mental saccade signal. This architecture, shown in Fig. 6, checks if advancement towards a goal is plausible using the perceived object. Such an assessment requires the prediction of what would happen if an action is performed that utilizes the perceived object. The prediction is based on the ML schema [29] and previously developed needs and learned approaches to their implementation.

In Fig. 6 let us assume that currently the most active primed memory neuron is neuron H (representing my friend's house) and it puts his house representation in the working memory spotlight. Based on the existing system of learned actions and abstract pains developed through ML, this perceived object will be associated with the intended action that results in an expected beneficial change in the environment. This is represented in Fig. 6 by the activation of neuron n. For instance H can be associated with organizing a garden party we planned long time ago at his house ($A_2$), and by delivering on our promise we would satisfy one of our needs to maintain friendship. Thus, the expected beneficial result of action $A_2$ would satisfy our need by increasing friendship, represented by n, and lowering the pain of not having n. As in ML, the perceived perception of a resource (represented here by friendship) lowers the pain related to an insufficient amount of this resource via an inhibitory link between n and $P_n$. This lower pain would be detected by a potential pain reduction circuit that would inhibit the next mental saccade through an inhibitory link.

The next mental saccade signal will be triggered, once a new winner of the WTA competition in the primed memory is established. Output of the "new winner established" block will be generated with a delay sufficient for evaluation of the intended action potential to reduce the pain signal that could satisfy a need that machine has. Notice, that pain signals are generated internally by the machine to reflect various goals and are dynamically changing with the perceived changes in the environment (goal completion) and the internal state of the machine (e.g. temperature or energy level).

For instance, if we recently had a party in my home, our friendship was reestablished and the corresponding pain signal $P_n$ would be low. Thus its potential reduction as an expected result of action $A_2$ may not be significant enough to trigger the pain reduction response and to block the next mental saccade. Thus, whether we think that organizing a garden party is a good idea or not depends on the internal state of the level of the pain signal related to our friendship.

A successful evaluation of the object in the attention spotlight will block the mental saccades from switching to another concept and allow memorization of the discovered solution in episodic or short term memory (learning block in Fig. 2) as well as planning for and/or executing an action.

Occasionally, the next mental saccade may be forced by the selection of a new motivation to act. This may result from the ongoing changes in internal pain signals and a significant increase in a pain signal may redirect machine to another action.

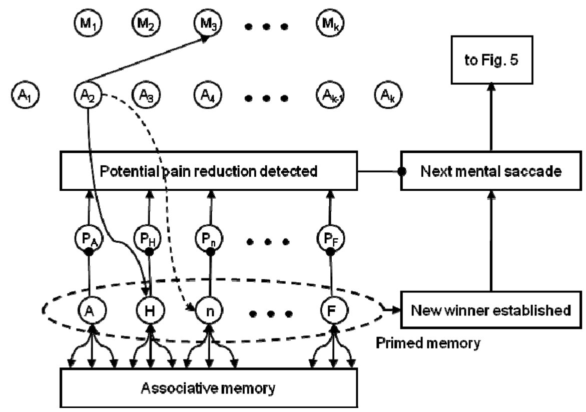


Fig. 6. Inhibition of the next mental saccade.

### *D. Action planning*

As an object in the attention spotlight was deemed useful to reduce a machine's pain, a reality check needs to be performed to see if the desired object is available in the environment or not. If the object is available and the machine makes a decision to perform the action it moves to action control. If the object is not available the machine may search for it or plan to get it. The search would set the machine's current goal to finding the object and either start the action of moving around the environment to find it or plan for where it can find it. The first one leads to a new goal and initiates search action control. The second one leads to further mental saccades related to the desired object. For instance, if we decided to search for sugar, and we cannot find any, we may decide to buy it. Notice that this is possible when next mental saccade is blocked and the object (sugar) remains in the focus of attention, allowing the activation of its associative memory (sugar and the ways to get it). This changes the objective to first find money and then to buy the sugar.

A more advanced plan for finding the object will involve several steps that need to be processed before the action control is triggered. If, for instance, the easiest way to get sugar is to buy it, we may mentally saccade to money and then realize that we do not have any cash. The next mental saccade is to go to the bank, to get some money. Thus a subgoal of getting sugar is created. This in turn switches on a subsequent subgoal, like getting into the car and driving to the bank. Each of these steps is either completed mentally (for instance we can remember that we have sufficient amount of cash in our pocket) or requires a search action (for instance checking the cash in the pocket). Thus planning and search actions are performed sequentially, until the overall plan is completed and we can start on the major action control to bring the sugar.

These mental searches and setting of subgoals can be performed within the cognitive structure illustrated in Fig. 7. In Fig. 7 there are a few important additions to the cognitive architecture from Fig. 6 that are needed to distinguish perceived reality from inner thoughts. First, the primed memory neurons are duplicated by creating a dual memory layer, and dual memory neurons take the place of primed memory neurons in competition for the cognitive attention spotlight. The second change is the duplication of the pain signals and dual pain signals, which are evaluated for potential reduction of the pain signal. The third change is the addition of the intention layer that contains neurons $I_m$ that represent the intended actions $A_m$, m=1,…,k.

A dual memory neuron in the attention spotlight is evaluated to see if it can be used to reduce the pain signal. Suppose that as before neuron H (representing sugar) won the attention spotlight competition and it was determined that by performing action $A_2$ (add sugar to the tea) the pain $P'_n$ (need to make tea) will be reduced. Now let us assume that sugar was not really observed (or easily found in the view) and we need to plan to get it. So, although sugar is required no real sugar is perceived and H is not active. Thus action $A_2$ will be blocked since the necessary ingredient, sugar, is not observed. This can be easily done by providing a disinhibitory link to $A_2$ from H (not shown). Thus, the system determined its current need to provide resource H by activating the pain signal $P_H$. Signal $P_H$ activated intended action $I_1$ that, as the system learned from ML, is required to obtain H. In this case it will be the intended action of buying sugar. This can only be provided if money is perceived in the attention spotlight. If K stands for money, then $P_K$ and subsequently $P'_K$ will be activated to indicate the system's interest in finding money. A mental saccade to signal K' will put money in the spotlight and if it can be observed, then desired action $A_1$ will be performed, resulting in the appearance of H that can be used to perform the original task of adding sugar to tea. However, if money cannot be found $P_k$ blocks $I_1$ and the system will need to perform action $I_4$ to get money from the bank. If more steps are required, deeper subgoal structures can be used as discussed in [29].

Notice that as soon as intended action $I_2$ was activated it activated the pain $P_H$, and $P_H$ blocked the intended action. The intended action was reactivated as soon as H was observed and blocked the activation of $P_H$. Several activated pain signals can be present simultaneously, allowing opportunistic behavior. For instance, if the system searched for money and instead found sugar, it uses it to sweeten the tea by removing the inhibition of both $I_2$ and $A_2$. Pain signals are deactivated on the completion of a corresponding subgoal.

In the presented model there is one pain and dual pain neuron for each primed memory and dual memory neurons and they have predesigned fixed interconnection structures as shown on Fig. 6. Neurons I act as gates for neurons A, such that there is one I neuron for every A neuron. The system learns the connection strength between pain signals and actions that eliminate these pain signals. Since actions require resources to act there are fixed inhibitory and excitatory links between actions and required resources and between actions and motor control neurons M.

### *E. Action control*

The last major block of cognitive architecture presented in Fig. 2 is action control. While deciding if to perform an action, consequences of performing it must be considered.

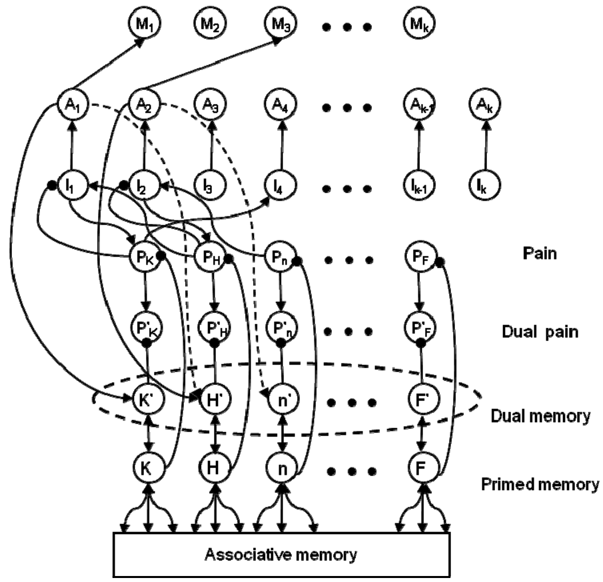


Fig. 7. Architecture to plan action with subgoals.

After an action plan is completed, it must be executed in the environment and environmental conditions are not always good to perform an action or consequences of the action may inflict a pain. For instance, whereas the normally acceptable solution is to go to store to buy sugar, the store may be closed, or the road impassable due to flood water. However, if there is no predictable negative consequence, the action may begin.

A block level action control schema is shown on Fig. 8.

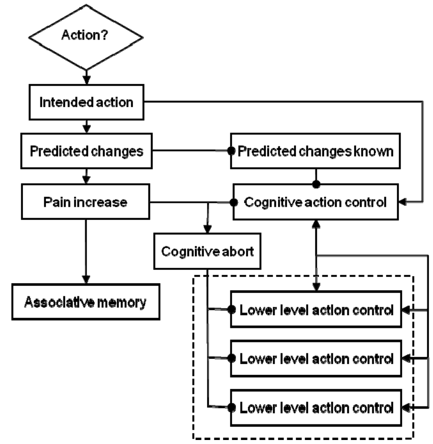


Fig. 8. Action control block

Well rehearsed actions in a familiar and unchanging environment require very little cognitive attention and can be preformed according to memorized sequences of motor control functions. Once a decision to perform this kind of action is made, the action is monitored occasionally by the cognitive center, while controlled with a large degree of automation by lower level action control (LLAC) sequential memories. Several LLACs can be simultaneously active. For instance, we can walk, talk to a friend, and carry a tray with a hot coffee in a cafeteria. When we suddenly see a threatening situation (a bus boy pushing a cart in front of us) we can simultaneously warn a friend, halt our walk, and protect the coffee from being spilled. This requires a cognitive intervention in 3 LLACs – issue a warning, stop walking, and balance the tray. Since this intervention requires proper motor coordination (like changing pitch and volume of our voice, increased muscle strength to halt our walk, and proper balancing of a tray that depends on the rate of deceleration) this coordinated breaking activity may need to be trained, since it has to be almost automatic.

There are several levels of action control. On the lowest level of reactive response, action control is almost automatic and does not involve cognitive decision making. On the lower cognitive level we simply evaluate if the dominating pain that triggered the action was reduced or not (self-centered approach), but on the highest socially acceptable action control we consider higher order pains, that require proper training and advanced social interactions to develop.

For instance if we are hunting and see a deer in the shooting range, we could execute the shot to satisfy our need. However, if a young fawn shows up before we pull the trigger, we may reconsider, feeling a pain for the child left without a mother to feed or protect it.

The proposed model of action control combines grounded object representation developed by the blackboard approach and abstract pain system developed in ML.

## VI. CONCLUSION

This paper presents the computational model of a new cognitive architecture and its attention switching mechanism based on internal motivations and mental saccades. It is intended to control behavior of robots in real environments and avatars that need to learn how to operate in virtual worlds. A machine's actions are directed by pain driven internal motivations to explore the environment and to accomplish its goals. Mental saccades provide a mechanism that moves an attention spotlight sequentially to primed associative memory areas for evaluation. At the same time, it is the mechanism behind the visual saccades, when new attention spotlight is driven by the visual input. Visual saccades and other forms of perceptions, together with internal motivations and cognitive processes, switch the machine's attention. Competition among the machine's attention signals is subconscious and only the winner signals are recognized cognitively. By responding to changes in the environment, the machine can redefine its goals and take advantage of new opportunities.

## REFERENCES


[1] F. van der Velde and M. de Kamps, "Neural Blackboard Architectures of Combinatorial Structures in Cognition," *Behavioral and Brain Sciences* 29 (1):37-70, 2006.

[2] J. Laird, P. Rosenbloom, A. Newell, Soar: an architecture for general intelligence, *Artificial Intelligence* 33 1987.

[3] J.R. Anderson, C. Lebiere, The newell test for a theory of cognition, *Behavioral and Brain Sciences* 26, 2003.

[4] S. Shapiro, et al., Metacognition in sneps, AI *Magazine* 28 2007.

[5] P. Langley, An adaptive architecture for physical agents, in: Proceedings of the *2005 IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, 2005.

[6] J. Hawkins, S. Blakeslee, *On intelligence*, Brown Walker (2006).

[7] R.C. O'Reilly, T.S. Braver, J.D. Cohen, A biologically-based computational model of working memory, in: A. Miyake, P. Shah (Eds.), *Models of Working Memory,* 1999, pp. 375–411.

[8] R. Hecht-Nielsen, *Confabulation Theory: The Mechanism of Thought*. Springer 2007.

[9] G.M. Edelman, Neural Darwinism: Selection and reentrant signaling in higher brain function. *Neuron* 10 (1993) 115-125.

[10] R. Sun, X. Zhang, Top-down versus bottom-up learning in cognitive skill acquisition. *Cognitive Systems Research* 5, 63-89, 2004.

[11] A. Nestor, B. Kokinov, Towards Active Vision in the DUAL Cognitive Architecture. *International Journal on Information Theories and Applications* 11, 9-15, 2004.

[12] N.L. Cassimatis, Adaptive Algorithmic Hybrids for Human-Level Artificial Intelligence. *Advances in Artificial General Intelligence.* IOS Press. Eds. B. Goertzel and P. Wang, 2007.

[13] B. Goertzel, Opencog prime: a cognitive synergy based architecture for embodied artificial general intelligence, in: *Proceedings of the ICCI-09*, Hong Kong, 2009.

[14] B. Goertzel, et al., World survey of artificial brains, Part II: Biologically inspired cognitive architectures, Neurocomputing (2010), doi:10.1016/j.neucom.2010.08.012.

[15] S. Franklin, The LIDA architecture: Adding new modes of learning to an intelligent, autonomous, software agent. In *Proc. of the Int. Conf. on Integrated Design and Process Technology*. San Diego, CA: Society for Design and Process Science, 2006.

[16] J. Bach, *Principles of Synthetic Intelligence*, Oxford University Press, 2009.

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

[18] A.D. Baddeley, "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., 1993.

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

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

[21] R. Desimone and J. Duncan, "Neural mechanisms of selective visual attention," *Annual Review of Neuroscience*, 18, 193-222, 1995.

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

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

[24] J.A. Starzyk, "Motivated Learning for Computational Intelligence," in *Computational Modeling and Simulation of Intellect: Current State and Future Perspectives*, edited by B. Igelnik, IGI Publishing, 2011.

[25] F. van der Velde, and de M. Kamps, "From knowing what to knowing where" *J. Cognitive Neuroscience* 13: pp. 479–91, 2001.

[26] C. Koch and T. Poggio, "Predicting the visual world: silence is golden," *Nature Neuroscience*, vol.2, no. 1, 9-10, 1999.

[27] J.P. Gottlieb, M. Kusunoki, and M.E. Goldberg, "The representation of visual salience in monkey parietal cortex," Nature: 481–84, 1998.

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

[29] J.A. Starzyk, J.T. Graham, P. Raif, and A-H. Tan "Motivated Learning for Autonomous Robots Development", *Cognitive Systems Research*, Available online 12 Jan. 2011.