A Biopsychically Inspired Cognitive System for Intelligent Agents in Aerospace Applications

Journal:	*Infotech@Aerospace 2012 Conference*
Manuscript ID:	Draft
luMeetingID:	2607
Date Submitted by the Author:	n/a
Contact Author:	Xu, Xudan

A Biopsychically Inspired Cognitive System for Intelligent Agents in Aerospace Applications

Xudan Xu[1]
Beihang University, Beijing 100191, China
Ohio University, Athens, OH, 45701, USA.

James Graham, Basawaraj[2,]
J. Jim Zhu, Janusz A. Starzyk [3]

Ohio University, Athens, OH, 45701, USA.

Ping Zhang[4]
Beihang University, Beijing 100191, China

The current state-of-the-art machine intelligence cannot meet the challenges of uncertainty, complexity, time urgency and rapidly changing nature of the dynamic environment in many demanding aerospace applications, such as autonomous extraterrestrial exploration. A biopsychically inspired cognitive system is proposed based on the biological Central Nervous System architecture. The main cognitive units, Decision Making and Planning, Motivation & Learning and Long-term Memory are explained in detail. The algorithms derived from the Central Nervous System and former works on motivated learning are shown for motivation, memory, attention switching, decision making, and planning together with obtained simulation results. The proposed cognitive system can be used for intelligent agents for autonomous scientific exploration, as well as in agriculture, rescue and military applications or as a basis for cooperative groups.

Nomenclature

P	pain or need;
B	pain stimuli array;
R_K	known resource state;
R_P	preferred resource state;
P_{sat}	primitive pain saturation level;
r_{p}	primitive pain rate constant;
τ_{p}	primitive pain time constant;
A	Action;
O	Object;
U	Utility for each task;
ML	motivated learning;
RL	reinforcement learning;
w_{BP}	pain stimulus weight;
n	elapsed (discrete) time since last reset of primitive pain;
Δp	pain reduction by task
Δc	cost resource consumption by task
T_{fin}	percentage completion by task
ΔO_i	decrement of weight for the *i*th Pain-Object links for non-winning goals;
ΔA_i	decrement of weight for the *i*th Pain-Action links for non-winning goals;
μ_O	maximum increment of the Pain -Object link weight
μ_A	maximum increment of the Pain -Action link weight
α_O	upper limit for the Pain-Object link weight
α_A	upper limit for the Pain-Action link weight
w_{PO}	Pain-Object weight
w_{PA}	Pain-Action weight
w_C	task cost weight
w_P	task pain reduction weight
w_T	task completion weight

[1] Xudan Xu is a Ph. D. candidate at the Science and Technology on Aircraft Control Laboratory, Beihang University, Beijing 100191, China, and is currently a visiting Research Associate at the School of Electrical Engineering and Computer Science, Ohio University, Athens, OH, 45701, USA. Email: xx505111@ohio.edu. She is a student member.
[2] James Graham and Basawaraj are Ph. D. candidates at the School of Electrical Engineering and Computer Science.
[3] Janusz A. Starzyk and J. Jim Zhu are professors at the School of Electrical Engineering and Computer Science. And J. Jim Zhu is an Associate Fellow of AIAA
[4] Zhang Ping is a Professor at the Science and Technology on Aircraft Control Laboratory.

I. Introduction

Outer space has always been both a mystery and a challenge for human beings. Forty years ago, for the first time man began to leave their home planet and journey to the moon. Today, extraterrestrial exploration remains a quest for mankind. Spirit, Opportunity, Curiosity and other intelligent agents have been or will be launched into the space to help us learn more about extraterrestrial bodies. For agents exploring the outer space, the environment that they must cope with will be quite demanding, as a result of the high uncertainty, complexity, time urgency and rapidly changing nature of their missions. Moreover, complex missions for the agents cannot be completely planned a priori. Although the current state-of-the-art machine intelligences have made significant strides in recent decades, they still have not met these challenges.

Most of the current state-of-the-art machine intelligences are remotely commanded or human supervised, which require human intervention in situation awareness and assessment, goal setting, decision making and mission planning. This necessitates significant communication bandwidth and power, or else the agents will not be able to respond to dynamic environment timely and effectively. Moreover, at the present most of the practical machine intelligence are pre-programmed; machine knowledge are static; and machine learning are supervised. As such the agents are not able to set goals to pursue when facing complex operating scenarios, nor innovate or improvise in dealing with unknown environments or unforeseen situations autonomously. While supervised learning increases the onboard knowledge base, it does not increase the wisdom of the agent. The pre-programmed intelligence also lacks distinct cognitive traits, or personalities, to facilitate complementary or symbiotic autonomous collaborations. Comparing to the biological counter parts, the current agents are still too machine-like. Thus, a cognitive system that gives agents some intelligence and enables them to learn about their dynamic environment and make decisions cognitively is necessary[1-6]. In particular, some U.S. research funding agencies have recently identified the needs for developing intelligent agents "… with the capability for situation awareness, allows for greater degree of uncertainty in terms of reasoning systems and produces greater robustness and adaptability in planning algorithms in dealing with unexpected interruptions and rapidly changing objective."[7] The system should "not only ascertain their descriptive validity and neural plausibility or feasibility, but also deepen understanding of mathematical characterizations of principles of adaptive intelligence"[8].

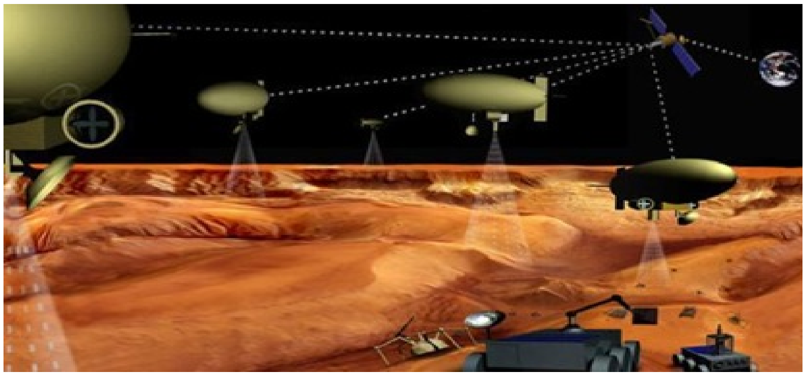

Figure 1. A team of agents performing an extraterrestrial exploration (Ref. 9)

In this paper, a biopsychically inspired cognitive system architecture for aerospace application is proposed. Compared with preprogrammed intelligence, the cognitive system with advanced machine intelligence should possess the following capabilities:

1) Self-motivated rather than directed goal setting and pursuit: advanced machine intelligence sets, pursues and modifies goals guided by the innate "pain and gain" sensation in order to survive and thrive in the unknown environment;
2) Evolved rather than programmed cognition: the programmed cognition is "if-then" cognition, but the advanced machine cognition driven by the need to survive and thrive allows the intelligence to evolve from the primitive reflexive response to environment stimuli to more advanced deductive causal analysis, inference, improvisation and innovation, leading to abstract goal setting and pursuit;
3) Capable of unsupervised learning: learning is driven by a need and is accomplished by retaining the ability and knowledge that satisfies the need. The motivated goal setting enables unsupervised learning based on the built-in reward-penalty mechanism;
4) Able to improvise, innovate and imagine: one of the genetically encoded primitive pains for animals and humans is the anxiety facing an unknown environment or situation. Exploration and trial-and-error, which are the primitive "pain-relievers" endowed by nature to all spices that have survived the evolution, naturally lead to improvisation, innovation and imagination capability;

5) Possess distinct personality and unique abilities: as different persons have different personalities and capabilities, with different structural and algorithmic parameters and learning history, the cognitive system should endow the agents with different personalities and capabilities, which will provide the agent community with more collective capabilities and robustness for surviving and thriving in the unknown environment.
6) Capable of cooperating: sharing data obtained from the environment, negotiating to make a decision and executing the decision for mutual benefits.

Extraterrestrial exploration and colonization will entail groups of large number of heterogeneous agents in collaboration, such as rovers and robots on the ground, flying disks or blimps in the air and orbiters in the orbit, as depicted in Fig. 1. In this paper, we study how agents can adapt to information-rich environments that are uncertain, dynamically changing and often adversarial in nature, and gain knowledge and expertise to make decisions with effectiveness. An architecture for the cognitive system, inspired by the biological Central Nervous System (CNS), is proposed in Section III. A biopsychically inspired motivated-learning cognitive algorithm is also proposed in Section IV. The simulation results are reported in Section V, and Section VI presents the conclusion of our work.

II. Central Nervous System

An understanding of the Central Nervous System (CNS) is inspiring and motivating for the development of cognitive systems, especially biopsychically inspired systems such as those proposed here. Here we provide a brief overview of the current understanding of the CNS, mainly based on Refs. 10 and 11. The interested readers are advised to consult these and other references for more details.

The CNS is the most complex structure in nature, producing the phenomenon of conscious behavior and imagination. It has been found that the CNS is hierarchical, consisting of the spinal cord, the brain stem and the forebrain. The higher level has more general concerns. At the lowest level, the spinal cord, only a single group of muscles is concerned; at the slightly higher level, computations concern several muscle groups to perform a coordinated action; while at the highest level, the forebrain, the whole body is coordinated in planning and execution of a goal oriented activities.

The spinal cord serves as a data transmission bus, relaying commands from the brain to muscles, and somatic sensory signals from muscles and skin to the brain. In addition it coordinates some reflexes. There are three types of neuronal circuits in the reflex system: 1) diverging circuits to spread the signals to the muscles; 2) reciprocal inhibitory circuits to relax the muscles and 3) oscillatory circuits to prolong the withdrawal action. It is also learnt that, there are three types of motion: 1) the reflex motion; 2) the voluntary motion, and 3) the rhythmic motion. The reflex motion is the basic motion and can hardly be influenced by high level consciousness. The voluntary motion is a goal-directed motion, which is controlled by high level consciousness. With enough practice voluntary motion will be learnt and stored in the procedural memory, commonly known as muscle memory, thereby becomes rhythmic motion. For example, playing the piano is highly reliant on muscle memory. The rhythmic motion is a combination of the reflex and voluntary motions. It is started or stopped by the high level consciousness, with most of the process being automatic, without any high level conscious guidance, similar to walking or riding a bicycle.

The brain stem includes several segments such as the medulla, the reticular formation and the midbrain. The medulla is responsible for body balance and equilibrium. The reticular formation performs low-level motion and the midbrain is involved in vision and sound localization. The brain stem is typically the highest level of computation for low-level vertebrates. Its functions are mainly enabling motion.

The forebrain is found in highly evolved vertebrates. It contains the highest level computational center, the cerebral cortex, and other parts, such as the cerebellum, the thalamus, the basal ganglia and the limbic system. The cerebellum plays an important role in motion control, especially the coordination control. The thalamus is a significant computational and switching center through which all necessary information is routed before entering into the brain. The cerebral cortex, consisting of the frontal lobe, parietal lobe, temporal lobe, occipital lobe and limbic lobe, covers all high level CNS structures. The frontal lobe has functions related to the ability of sequentially organizing complex tasks, long-range planning and abstractions. The temporal lobe is also well known for processing auditory inputs. The occipital lobe is directly connected to the nerve bundles of the eye and deeply involved in visual processing. Its outputs go into the temporal lobe, parietal lobe and more than 30 other regions in the brain for further processing. The parietal lobe is a high level motion planning and monitoring center, which receives visual, auditory, vestibules and somatosensory signals, and directs arm, hand, and eye movement. The limbic lobe is the geographical seat of emotions and it is also involved in memory management. It contains the memory manager, hippocampus, which is important for the consolidation of information from short-term memory to long-term memory. The hippocampus is one of very few brain regions where new neurons continue to be created throughout life, so that the vertebrate can always learn new environments and perform spatial navigation[12] .

The algorithms in the CNS are hierarchical too. The primitive motivations, such as hunger, thirst, abnormal temperature, sickness, fear etc., are survival-driven. Every primitive motivation has a relevant “pain”, if the pain level is high, which means an agent is facing danger, the agent takes a series of actions to reduce the pain. The thalamus is the “pain center” in vertebrates. It receives all the sensory signals, calculates the pain level and sends them to the brain.

For a slightly abstract motivation, like boredom, the intelligence evolves from the primitive response to pain to a more advanced deductive causal analysis. For example, if a rat can obtain some food by pressing a button, it will keep pressing. But if penalty is obtained by pressing the button, the rat will not press it. Gradually, biological creatures form an adaptation mechanism by such "penalty - reward" learning.

Moreover, there is another "hormonal algorithm" in CNS for sensory-motor control system reconfiguration. Hormones are carried everywhere in the body by blood (as opposed to neural transmitters carried through the nervous paths) to change the sensitivity of sensory and musculoskeletal systems under emergency and stressful situations. This method provides a tradeoff between performance, energy consumption and robustness. In normal mode, the agents operate in an economical and robust manner. But when agitated, the agents need to generate a response in a fast and powerful way to achieve better performance, often at the cost of increased energy consumption and reduced robustness. For example, before a competition, an athlete is more excited and more active than usual because of increased level of adrenaline (epinephrine), adjusting the body for a competition. Such mechanism supplies safety and robustness.

With the principle of “survival of the fittest”, the structure and mechanism of the CNS has been naturally optimized. First, it is efficient. Every neuron is a computing device, operating on its own set of inputs and transmitting its outputs to a specific set of other computing devices. Many different computations are performed simultaneously in many different units. Second, it is robust. Each device performs only one task at one time. No part is very complex. It is a network of computing devices, instead of a single processor. If one single part is damaged, its function can be taken by another part. Third, it is massively parallel. It takes advantage of a network of billions of simple and slow computing units, operating in parallel to provide high computation capability and fast computation speed. These salient features should be used in guiding advanced machine intelligence architecture and algorithm design.

III. The Proposed Cognitive System Architecture

The proposed cognitive system for intelligent agents, shown in Fig. 2, is inspired by the biological CNS. The bottom-left part of Fig. 2 describes the sensory inputs of agents, which sample environment information, such as vision, sound, taste, touch and smell. These signals are sent to the sensory processor, which filters the input sensory data, changes it into the abstract format for higher level usage and enhances those that are of interest according to commands given by higher level computing centers, such as the Decision Making and Planning (DM&P) unit and the Motion Monitoring and Coordination (MM&C) unit.

The DM&P is the intelligence center of the cognitive system. When making a decision, the DM&P unit accepts the data from sensory processor and queries the long-term memory for related information. All these data are loaded into the working memory to play out various scenarios by mental simulation to predict outcomes, assess benefits and costs as well as consequences. After making a plan, the DM&P unit sends sequential organizing commands to the MM&C unit and writes the useful results into the long-term memory, which corresponds to "learning".

The Memory Manager plays the same role as "Hippocampus" does for creatures, to facilitate learning. It is a bridge between working memory and long-term memory, turning the useful transient memory into long-term memory. It also features the spatial navigation. Many neurons in it act as place cells and agents use it to map environment information[14].

The MM&C unit is the motion monitoring and coordination center. Just as in vertebrates, the MM&C unit controls the voluntary motion of agents, such as taking photos, examining a new object. By practicing more and more, agents write the successful motor commands into the procedural memory, thereby turning them into rhythmic motion sequences. The MM&C unit can start or stop rhythmic motions, thus relieving itself from mundane motion planning and control tasks.

Both the Motion Control (MC) unit and Reflex Action (RA) unit are motor control units. The commands to the MM&C unit are transmitted to desired control reference commands for the MC unit, which then delivers final muscle control commands. The RA unit performs basic reflex motions for agents (such as the stretch reflex), mainly as spontaneous protective response, which does not need guidance from higher level consciousness, but can be inhibited by it if necessary, as in the case when the agent must endure some local physical damage in order to save its "life".

The motors of the agent execute motion commands and influence the environment with their action. At the same time, other agents affect the environment too. The agent feels all of these changes through sensors, thereby closing the sensory-motor control loop. Successful actions will be promoted to the long-term memories through the internal reward mechanisms, while actions that produced undesired consequences will be gradually degraded and eventually eliminated from the long-term memory by the internal penalty-reward mechanism.

For autonomous agents, the Health Management (HM) unit is a necessary segment, working in parallel with the CNS to regulate the agent's bodily functions and monitor the agent's health. The agent's health signals go to the HM unit, which takes care of primitive needs, such as the battery charge level, the internal temperature and pressure based on the internal sensors, so that it allows the agents to form their own directives or motivations based on the needs that it determines for itself (Primitive Pain), just as what the Autonomic Nervous System does for human beings. Motivation is primarily an intrinsic functionality, because competing processes within the agent determine its motivations. There is an extrinsic influence on how the initial primitive pains are endowed by the environment and evolution, but how the primitive and higher-level abstract pains are resolved is solely the purview of the agent itself and shaped by its particular environment and unique life experience. Because of its ability to determine its own pains and goals, a motivation based agent is capable of dynamic unsupervised learning. A motivated agent is therefore, more directed and more capable of operating in diverse environments where it has little or no supervision compared to a typical reinforcement learning agent[17] .

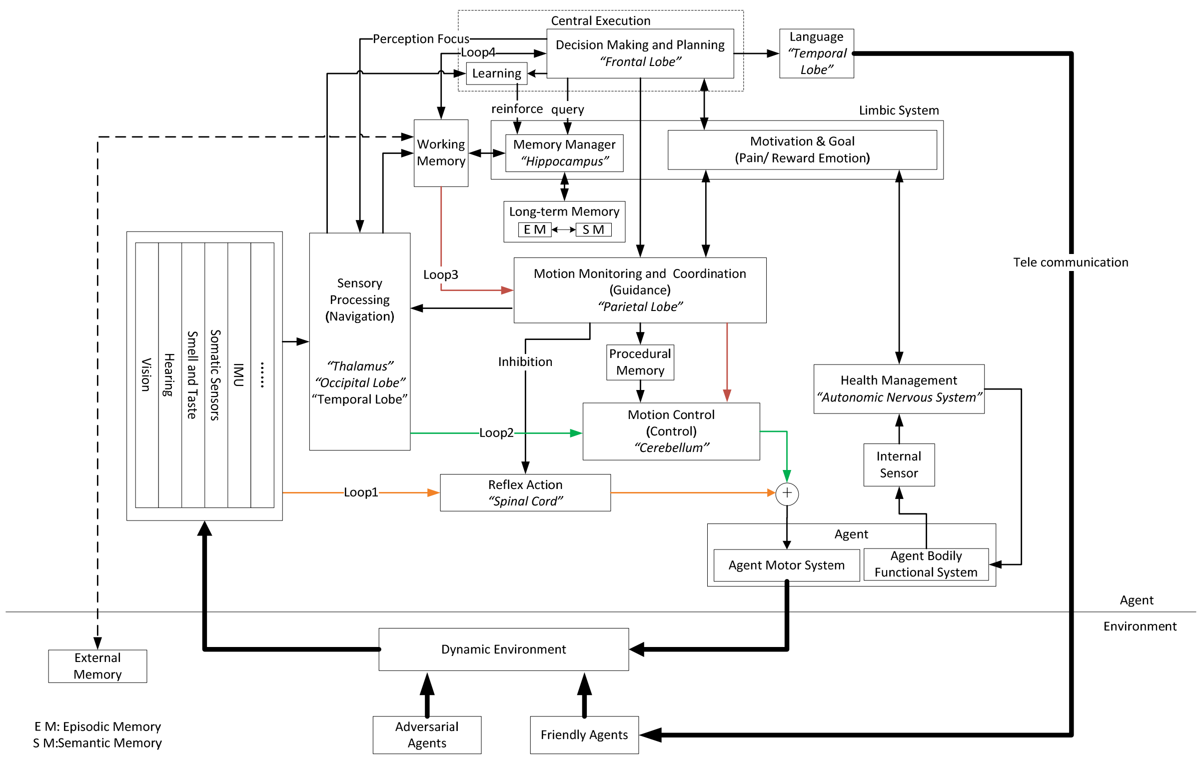

Figure 2. The cognitive system architecture

The memory acts as the central storage of all information, and different units query the memory for the information about the objects in the visual field and their relation to the pains, goals and motivations in the cognitive system. The memory provides the past experience with the different actions that can be taken on an object(s) in the visual field and the cost and benefit of taking those actions. It can also provide similar details about the objects that are presently not in the visual field. Memory has been categorized into various forms in literature; like episodic and semantic memory[21]. Semantic memory stores the generalized understanding of the world, though the understanding can be based on personal experiences. The semantic memory does not store any specific event. Thus the semantic memory is central to our understanding of the world and it gives meaning to the information stored in other memories such as working, procedural, episodic etc. All these memories are like database to agents. For decision simulation or planning, a form of "working memory" is also needed.

IV. The Key Algorithms for the Cognitive Process

In this section, three key algorithms are presented for the three high level cognitive units: DM&P Unit, Long-term Memory Unit and Motivation & Learning (M&L) Unit, which are implemented in a simplified version of the cognitive system proposed above, as shown in Fig. 3. The DM&P unit contains Decision Making block, Planning block and Attention Switch block. Long-term Memory Unit is divided into Semantic Memory and Procedural Memory. While Learning block, Motivation block and Primitive Pain block comprise the M&L Unit. The system is divided into three abstraction levels. The lowest level contains the Environment Model and Primitive Pain Model, which updates the environment information and agent pain level in real time; while at the highest level, the Decision Making block plans and makes action decisions and responds to important changes sensed by the agent.

As human beings have different personalities, the algorithm features different characteristic parameters, which can be initialized by the designer, and subsequently modified by the agents' experience through learning as desired. For different parameters, the results are different.

Motivated by the "hormone signals" present in vertebrates, the power supply network can be used to send "hormone signals", for electrically powered agents. They can be treated as global "carriers" for signals with different frequencies and values. The description of the main cognitive units is given below.

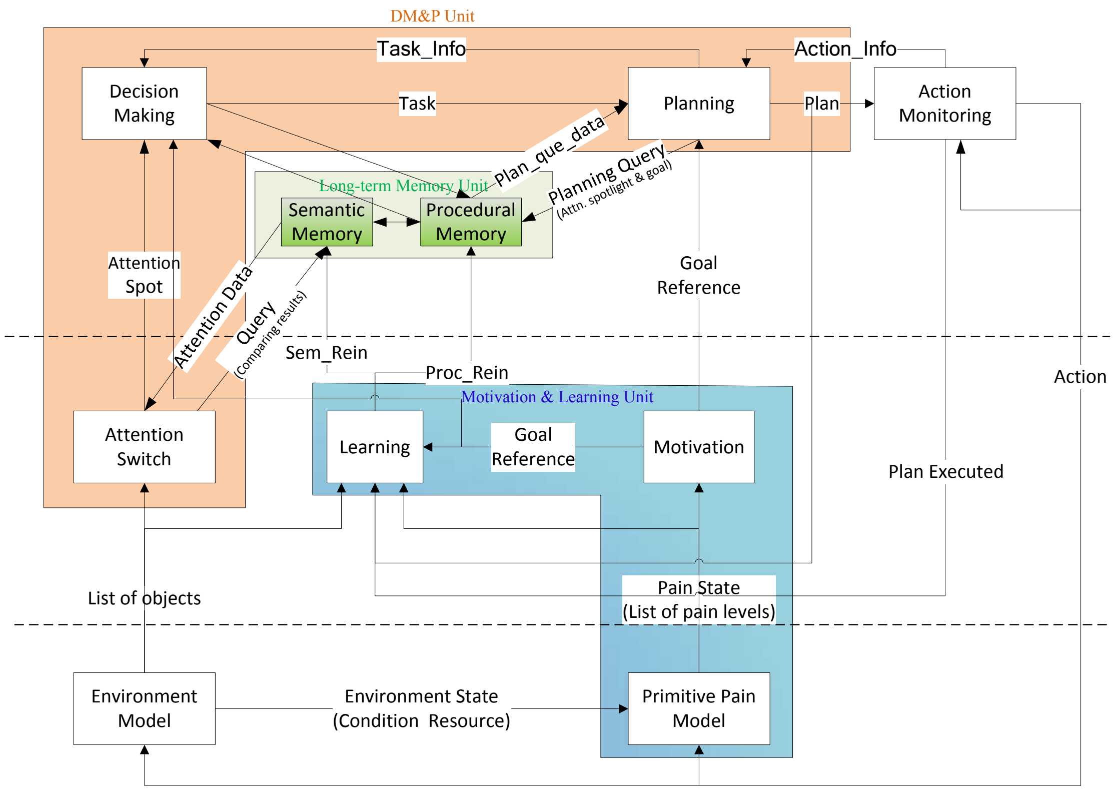

Figure 3. The algorithm mechanism for CNS

A. Motivation & Learning

This unit contains the Primitive Pain Model, Motivation block and Learning block. The Primitive Pain Model describes the basic needs and pains of the agent. The Motivation block derives the abstract motivations from the primitive ones and the agent's learning process. The learning behavior in agents follows Motivated Learning. The idea behind Motivated Learning begins with the concept of "embodied cognition" as defined by Brooks and Pfeifer[22,23]. The theory of Embodied Cognition assumes that a machine's "intelligence" develops according to its physical embodiment. The embodiment shapes and controls how the machine is able to interact and, thus, has a direct impact on how it develops. Motivated Learning subscribes to this principle in that it relies on the agent's interactions with the environment to set its needs and grow its motivations. Ref. 24 discusses in detail how motivated learning relates to the development of intelligence in an embodied agent. In comparison to reinforcement learning, a motivated learning (ML) agent has multiple value functions, sets its own objectives, solves the minimax problem, is stable, and acts when needed. In contrast, a reinforcement learning (RL) agent typically only has a single value function, relies only on externally set objectives, maximizes its reward (and is therefore unstable), and is always active. Comparing these lists leads to the observation that an ML agent would be more efficient and capable in an unstructured, dynamic environment than a similarly configured RL agent.

A ML agent works by perceiving the environment, as well as any internal status signals and translates them into needs or "pains" which the agent must act upon. An ML agent begins with one or more "primitive pains" that define its most basic needs. These primitives are relatively simple and are directly related to the functioning of the agent. For example, in humans, the basic primitive needs could be reduced to food, shelter, and reproduction. In artificial agents, the primitive needs would be similar, such as energy, operating temperature, and maintenance.

At the beginning of a simulation, an ML agent works like a reinforcement learning (RL) agent, since its needs map directly to the environment. While ML may have similarities to RL, its ability to derive it own motivations puts it ahead of many other learning methods as discussed in Ref. 17. In both ML and RL, satisfying basic needs alters the environment and consumes resources, however, in an ML agent it gives rise to new, learned, needs. For example, if an agent learns that it can "eat" an energy source at the expense of depleting the energy source, it will develop an "abstract need" for refilling the energy source and, once it is sufficiently depleted, devote some portion of its time to replenish the energy source. While an ML agent does not explicitly rely on curiosity, it can still use it once it has no other pressing needs to pursue; and does so by exploring the environment. Additionally, the agent can learn even from "incorrect" actions by keeping track of how its actions affect resources within the environment, even those it doesn't consider relevant.

In current implementations, the ML agent calculates its needs based on a simple formulation. It takes the "known" environment state R_K, and the preferred state, R_P, of the environment and compares them. The difference in the known and preferred state is then translated into a "pain stimulus" array, B, for the agent. The pains, or needs, are then calculated by multiplying the stimulus against their associated pain stimulus weights, w_{BP}. This is shown in Eqs. (1) and (2), respectively. The pain stimulus weights are initially zero, meaning that as far as the agent is concerned, the associated pains/needs do not exist until the agent discovers that they are relevant to its needs. This is where the reinforcement block of ML comes into play. As the agent discovers the importance/use of a particular resource, it reinforces the weight associated with its need for that resource. Primitive needs, on the other hand, are calculated differently, since they are always "known" to the agent. Equation (3) provides the function that has been used to determine the primitive needs for the agent. This function was chosen for two reasons; one, it partially mirrors the logarithmic growth of the bias calculation, and two, it reaches its maximum at about the same point in time as we attempted to make the primitive and abstract pains grow at similar rates. A linearly growing primitive pain could end up dominating the agent, not allowing it to resolve the abstract needs that may be required for it to resolve its primitive needs.

$$B = [B_i] = -\log_2 \frac{R_{K_i}}{R_{P_i}} \tag{1}$$

$$P_{abstact_i} = w_{BP_i} B_i \tag{2}$$

$$P_{prim} = f(n) = \begin{cases} P_{sat} \dfrac{r_p n}{r_p n + \tau_p}; & \text{since reset} \\ 0; & \text{when reset by agent action} \end{cases} \tag{3}$$

Motivation works with other units of an agent to provide the needed drives. Its most direct impacts are to the memory and the attention switching blocks.

B. Long-term Memory

In this section we take a close look at the memory structure, its learning algorithm and salient features. We also discuss about the advantages and disadvantages of our method. Memory, in animals, acts as storage of knowledge acquired and important observations and experiences. Any plans made and actions performed are based on the memory, and the plans and actions are optimized based on new observations and experiences. Thus memory plays a very important role in all animals. The memory unit in our agent plays a role similar to the memory in a biological being; it acts as a storage of the knowledge that the agent has acquired based on its observations and experiences. This memory is used by units, such as the DM&P, to perform their actions. The past observations and experiences are used in formulating plans and actions; the resulting changes in the environment are used to optimize the memory for future use. In our agent the memory receives environment data, pain levels and motivation information from the M&L unit, control information from the planning DM&P unit, and status of action from the Action Monitoring block.

The memory is based on a neural-network approach that uses weighted links to store associations. Initially the connections between the nodes are random and the focus of attention in the visual field saccades from one object to another, but as the agent learns to handle its pains the saccades are primed by the usefulness of the objects towards the agent's goals. Increasing pains force the memory to go through possible actions starting from the action with the strongest weights. The memory can perform two types of saccades: visual and mental. During visual saccades the spotlight moves from the object with most salient features (or its parts) to the one with less salient features. Similarly during the mental saccades the spotlight moves from the object currently in focus (either sensory or mental), to other objects or actions associated in the memory with the one in focus. The saccades are controlled by the planning block that evaluates how useful the concept in focus is for its goals. The attention switching block can start a saccade, but attention spot on a specific pain-action in the working memory (implemented in the DM&P unit) determines a potential action. As the memory saccades, the previously scanned concepts and / or actions are inhibited until associated areas of memory are searched through. Note that in mental saccades potential reduction in actual pain levels is evaluated without inhibiting real pains. The saccades are based on the idea proposed by Starzyk et al[18]. While the potential inhibition of the pain memory traces is based on Ref. 25. For the memory block, motivations act as priming signals. They help to direct attention to memory areas that are related to the most active pains, so that when the DM&P unit requests information from the memory, it receives the most relevant information related to dealing with the dominant motivations.

Each time the agent acts, the inputs to the memory change. The changes can come from the environment, the pain levels, the memory association resulting from mental saccades, and the attention focus. An action that leads to a decrease of a pain (P) causes an increase in the weight of the connection between this pain neuron and the action (A) neuron. There is also an increase in weight of the connection between this pain neuron and the object (O) on which the action was performed. However, if the action performed leads to an increase of pain then the weights are decreased according to the mechanism proposed for adjusting pain-action weights in Refs. 17 and 24.

Initial weights between the P-A neurons and between P-O neurons are set randomly in the 0-α_A and 0-α_O interval respectively. Similarly the maximum upwards or downwards adjustment of P-A and P-O weights are limited to μ_A and μ_O respectively. Equations (4) and (5) show the computation of weight adjustments for interconnections between P-A and P-O neurons respectively.

$$\Delta w_{PA} = \mu_A \min\left(\left|\alpha_A - w_{PA}\right|, w_{PA}\right) \tag{4}$$

$$\Delta w_{PO} = \mu_O \min\left(\left|\alpha_O - w_{PO}\right|, w_{PO}\right) \tag{5}$$

In addition, the connections from this pain to other actions and other objects are changed by the amount computed using Eqs. (6) and (7), respectively.

$$\Delta A_i = \Delta w_{PA}.w_{PAi} / \sum w_{PAi} \tag{6}$$

$$\Delta O_i = \Delta w_{PO}.w_{POi} / \sum w_{POi} \tag{7}$$

where $i \neq$ *winning Action (or Object).*

The main issues with this implementation are: a) the memory stores all possible actions that the agent is capable of (no new actions are learnt); b) cost of performing an action is not considered (no cost benefit analysis is performed); c) objects in the visual field can be identified unambiguously (no symbol grounding); and d) actions are one step processes (multistep actions are not represented). Currently, we are in the process of fixing these issues.

C. Decision Making & Planning

The blocks Attention Switch, Decision Making and Planning constitute this unit. Attention Switch allows the agent to immediately and directly shift the agent's attention spot if there is a sufficiently important change in the environment. Decision Making 's main function is to form the current To Do List based on the previous To Do List and the current attention spot. Then it sends the current task to the Planning block, where the Action list for the task is planned. The algorithm diagram is shown in Fig. 4.

In the diagram, Attention Switch gets the descriptive information of objects in the environment, runs visual saccade in long-term memory and picks the current attention spot, which will either benefit the current agent's motivation or prevent the agent from danger. The Decision Making block receives the attention spot from Attention Switch, querying long-term memory with object ID and highest pain level to get the associated activation and the new task for the attention spot. If the new task for current attention spot is emergent-which means that the task needs to be carried out immediately or the agent will be in danger, such as running away from a fire, or if the task will help to reduce the current pain level, the new task is considered. And the evaluation is run for not only the current new task but also the previous tasks, to update To Do List; otherwise, the evaluation is run only for previous tasks. In the Evaluation function, procedural memory is queried again to obtain real time associated information for each task. And the utility of each task is computed based on changed pain level, cost and how much the task has been finished, as that in Eq. (8).

$$U_i = w_P \times \Delta p - w_C \times \Delta c + w_T \times T_{fin} \tag{8}$$

The Evaluation function also evaluates the time window for each task. The Prioritization function prioritizes the potential tasks based on the results from Evaluation function. It deletes the tasks whose time window has expired, places the emergent tasks at the top of the list, and prioritizes the tasks according to their utility, from high to low. After the current To Do List is finished, the Command function gives the task name and task order to the Planning block. The To Do List can be updated every sampling period, even during the execution of tasks. The Task Monitor function in the Planning block processes commands from the Decision Making block first. It contains four cases – "Relinquish", "Defer", "Continue" and "Resume". For "Relinquish", the Planning block deletes the current task's action list and produces another action list for the new task; for "Defer", the current action list is pushed into the action stack before new list is made for the new task. For the "Continue" case the Planning block continues the current commands, while in the "Resume" case the deferred list is popped out from the stack according to the task name. The action list is obtained by querying the procedural memory and the current action is sent to Action Monitor block.

As we can see, the DM&P unit queries memory for different information, either for tasks which can be taken for the current focus of attention or for information associated to the current task. So a saccade control center is needed in the DM&P unit to tell which kind of saccade is expected.

In DM&P algorithm, the associated information, such as ∆p, ∆c, time window, is evaluated in real time based on the current environment and pain, and the action list is calculated by running mental simulations on improvised strategies. For every chosen task, the Planning block performs mental simulations with the information queried from procedural memory, to find an action list, which will advance the agent's goals.

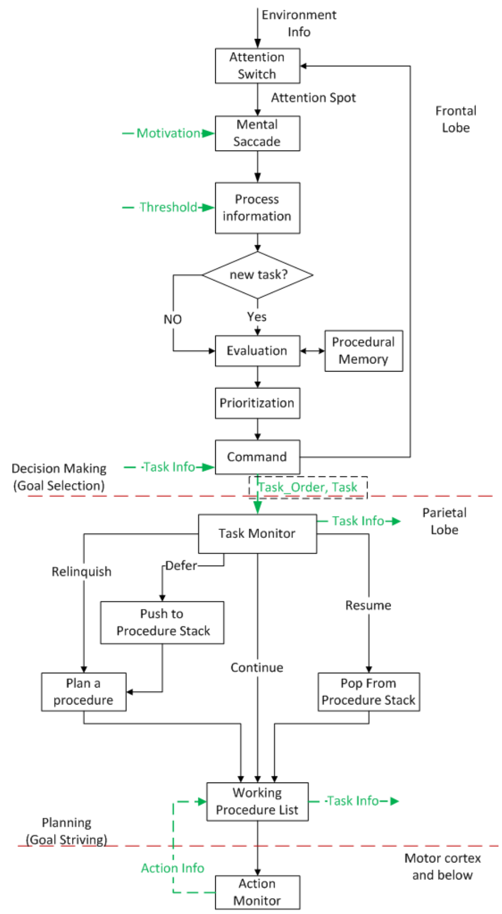

Figure 4. The diagram for DM&P

D. Others

The Environment model, Action Monitoring model and agent's action simulate the environment and the other parts of the agent except for the three units introduced above. The Environment model simulates the environment scenario for the agent. It interprets and responds to the agent's actions as needed, describing how the agent's actions influence the environment. And the Action Monitoring block monitors how the plan is executed, how far away it is from the goal and collects information for learning. The purpose of these units is to build a self-contained agent. While these units are essential for the operation of the agent, they are not currently our primary focus.

V. Simulation

A number of methods can be used to simulate an agent. First we developed a MATLAB implementation, ported it to C++, and are now in the process of integrating the code into the NeoAxis[26] game environment. NeoAxis is distributed as a free (for non-commercial use) game SDK. It contains a number of pre-existing resources and examples for environment creation, as well as both a map and a resource editor. Our MATLAB code allows easy debugging and concept development. However, it lacks a more complex environment and a good graphical presentation, hence, the reason for porting code to C++ and NeoAxis.

In MATLAB, we face several constraints, besides graphical limitations. The main constraint is the lack of a ready environment for the agent to operate in. We can create an agent as complex as we want, but without an equally complex and responsive environment to test it in, we are greatly limited. The current MATLAB implementation is limited to simple resource quantities and their respective locations, whereas a system like NeoAxis, provides the necessary underlying structures to design the environment and to interface it with our agent.

In this section, we present the results of simulations testing the main three cognitive units of the proposed agent. In the simulation, an extraterrestrial exploration scenario with the resources consisting of: Base, Energy pill, Sun, Shelter, Big rock, Medium rock and Small rock is assumed. The agent is equipped with a cooler and heater to adjust its temperature, and its actions are limited to 1) Use, 2) Go to, 3) Repair, 4) Turn on, 5) Turn off, 6) Break, 7) Build, and 8) Explore. The agent is deployed in an extraterrestrial environment with some basic knowledge in its memory and with 1) Damage, 2) Hunger, 3) Low Temperature, 4) High Temperature, 5) Mission, and 6) Curiosity as the primitive pains.

A. Motivation & Learning

Figure 5 shows a GUI we have developed for the MATLAB code to illustrate the agent's operation. In the center of the figure, we show the agent's environment. In this instance it depicts the surface of Mars (courtesy of NASA). Attached to this photo are several resources that have been "placed" in the environment. The agent itself is represented as a diamond inside square. As the agent moves in the environment and makes its decisions, its representation on the chart will indicate which resources it is using. Further, the internal pains (motivations) are represented in the lower bar chart arranged by the pain number. At the time the figure was taken the agent was only within the first three-thousand iterations, however, by this time it had already discovered how to resolve all of its primitive pains (pains 9-13) and has learned that at least three resources are of relevance to its needs (pains 1-3). For reference, the pain threshold for the agent is 0.3.

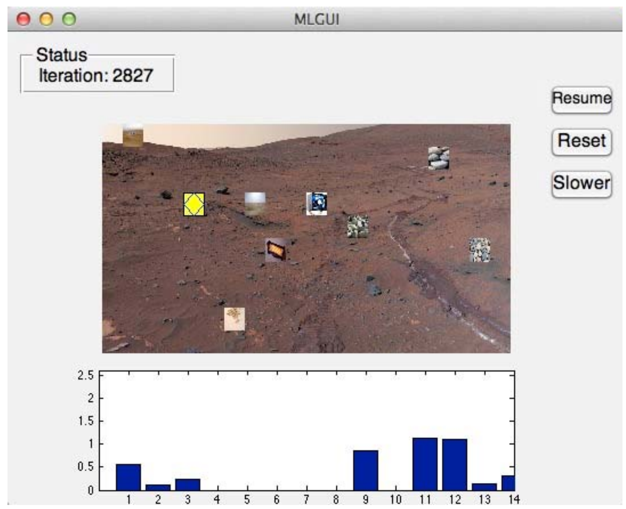

Figure 5. Motivated Learning GUI

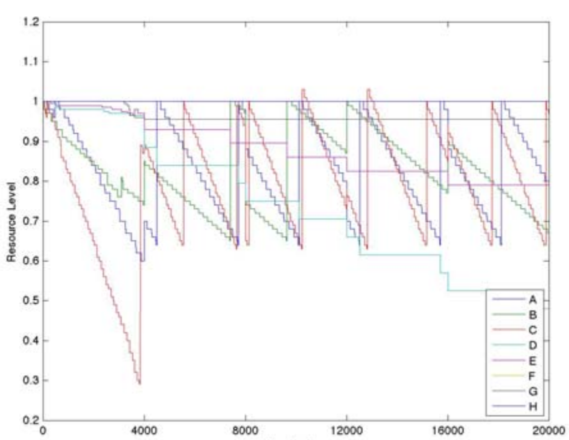

Figure 6. Normalized resource levels

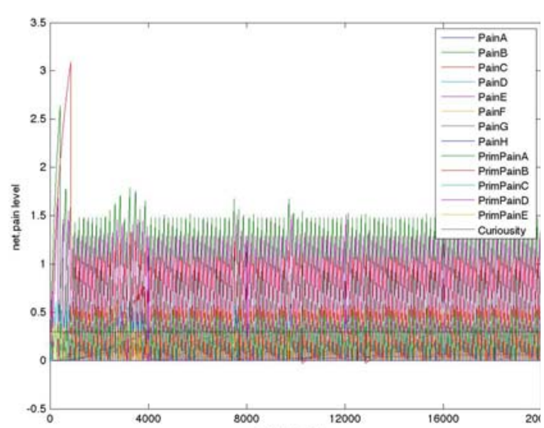

Figure 7. Pain levels

Figure 6 illustrates the resource usage of an agent running in the MATLAB environment simulation. As can be seen from the figure, despite an initial depletion of each resource, once the agent learns to use that resource, the resource level quickly stabilizes and reaches equilibrium. This is more apparent in Fig. 7, which shows various pain levels over time. In Fig. 7, distinct pain "spikes" are observed at times corresponding to either a new resource being learned or intense competition between multiple pains. Furthermore, after initial learning period, the agent has little problem operating from that point onward. As seen from Fig. 6, there are two resources that have yet to be restored by the agent (due to their associated pains not reaching a high enough level for the agent to act upon them), meaning that the agent may still need to learn more. However, based on previous runs, the agent should be perfectly capable of handling the situation.

B. Long-term Memory

Figure 8 shows another GUI in MATLAB. This illustrates the operation of the agent's memory, and shows the agent performing a visual saccade. The agent's environment is shown on the left and the object under the attention

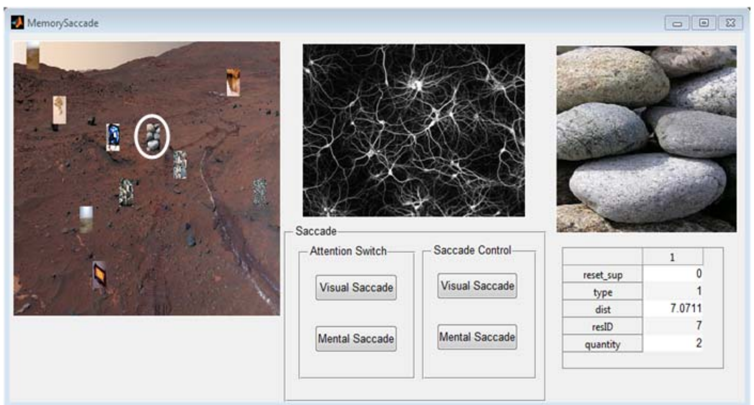

Figure 8. Result of Visual Saccade

Mental Associations	Definition
Pain	Low Supply-Small Rock
Goal	Break Medium Rock
Motor Action	Break
resID	Small Rock
resID	Energy Pill
Motor Action	Go to Base

Table 1. Associations for Mental Saccade

spotlight is highlighted, by a white circle drawn around it, and is shown on the right (medium rocks). A partial output of the memory is displayed near the lower right corner. The output consists of: type of saccade performed in this particular visual instance (type = 1); distance of the object from the agent (dist); name of the object (resID = 7, which represents medium size rocks); and quantity of the object present. This image was taken when the agent, running low on its supply of small rocks, comes across medium rocks and Table 1 shows the various pains, goals, objects, etc. that are associated with the object currently under attention spotlight (medium rocks). If the agent performs mental saccade related to the medium rocks, it knows that there is a pain, due to low supply of small rock. An associated mental saccade reveals, that the agent can solve this pain by breaking medium rock using the "break" motor action. Continuing this process of mental saccades, the agent knows that it will need "energy pills" to break medium rock and that it can find them at its "base". When the agent performs a mental saccade on the object under the attention spotlight it can either saccade through associated objects (or actions, etc.) only or saccade through other objects (or actions) associated with the object in focus as in the example shown above. Both mental and visual saccades are controlled by the decision making & planning unit.

C. Decision Making &Planning

The simulation results for the DM&P unit are shown in Fig. 9. The lower parts are the input from the environment (Panel_Resource, bottom right) and the current level of agent's motivation (Panel_Motivation, bottom left) separately. The environment information (resource) is given to the Attention Switching block to choose the current attention spot. The motivation is passed to the Decision Making block to decide the current To Do List, which will advance the agent's motivation. The To Do List is shown at the upper right part, which is obtained based on the evaluation results for the Potential List. At the upper-right corner are the task order indicators, to show which task order is sent by Decision Making block to Planning block. When we click on each task in the To Do List, the associated cost of resource and expected change in the pain level are shown in the middle graphs. The motivation graph, on the upper left, indicates the current motivation value arranged by various pains. Comparing the To Do List and the Potential List, we can tell the decision is made following the algorithm introduced in Section V: deleting the tasks that are no longer needed any more, and then placing the tasks in order, first from emergent task to nonemergent task, then from task with high utility to task with low utility. The steps for the current task are listed in the Action List.

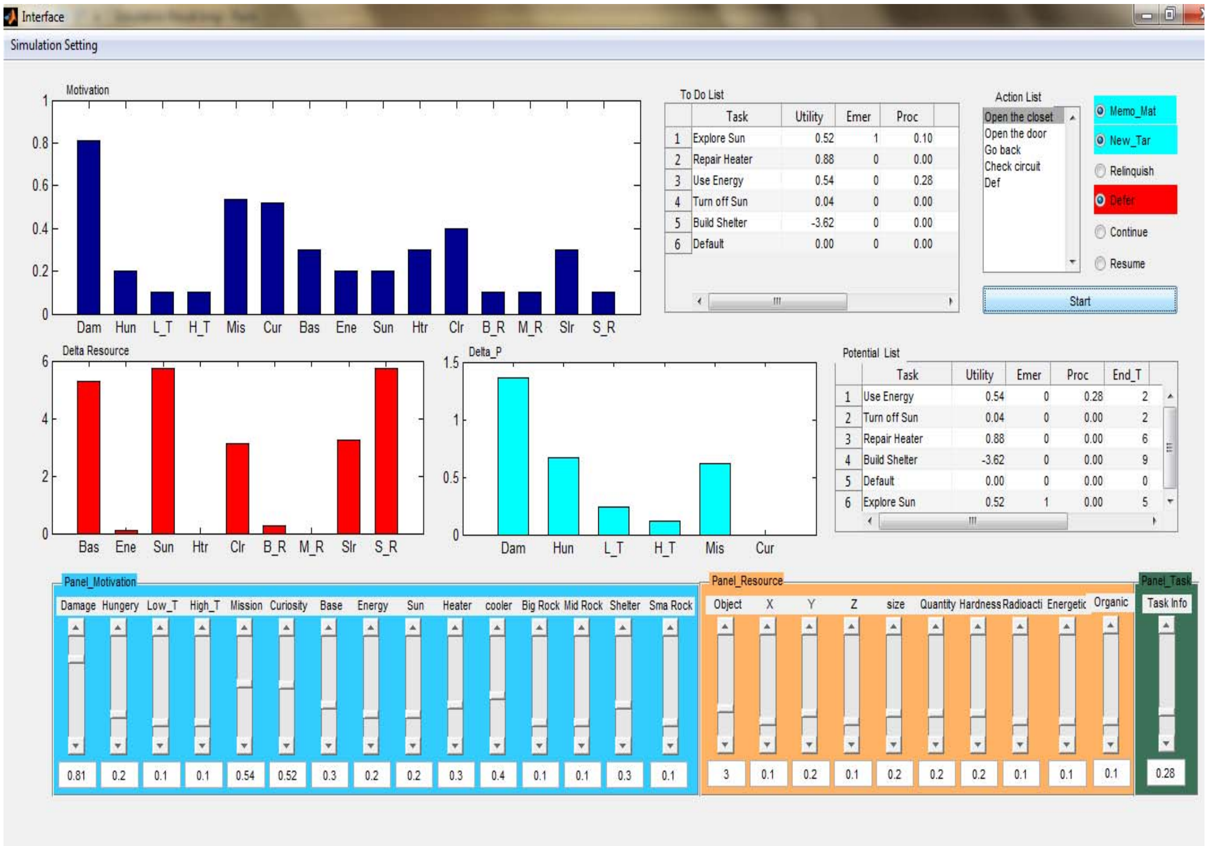

Figure 9. Result of DM&P

VI. Conclusion

In this paper, a biopsychically inspired advanced cognitive system is proposed for intelligent agents in aerospace applications. The proposed system is different from current intelligent agents such as Unmanned Aerial Systems (UAS) in that its cognitive intelligence is not preprogrammed, but rather it is based on survival-driven motivated learning that allows the machine intelligence to evolve from primitive reflex response to environmental stimuli to more advanced deductive causal reasoning, improvisation and innovation, leading to abstract goal setting and pursuit. It can learn the environment unsupervised and set the goals for its actions autonomously. This capability will enable the agents to operate in uncertain and demanding environment. With the principle of "survival of the fittest", the architecture and algorithms of the biological CNS are naturally optimal and allow new functions and capabilities to evolve. The modularity of the cognitive system is a network of computing units working concurrently. Each functional unit performs only one single task and all these units operate simultaneously to provide powerful computational capability. With such parallel and distributed network architecture, the system is readily implemented on microchips. Because many functional units can perform similar functions, the system features strong robustness. The simulation results of three main cognitive units for an extraterrestrial exploration scenario are presented, which verifies the proposed cognitive architecture and algorithms, The results are applicable to other aerospace application domains such as unmanned military missions, terrestrial unmanned scientific explorations, and unmanned search and rescue in battlefield and disaster aftermath[27,28]. In the future, further integration work is expected and the code will be ported into C++ and NeoAxis for more robust simulation and testing. The decision making and planning algorithm will be further developed based on human psychology.

Acknowledgments

Authors are grateful to Dr. Jeff Vancouver a professor of Psychology at Ohio University for insightful discussions.

References

[1]H. Liu, D. Gu, R. J. Howlett, Y. Liu. *Robot Intelligence-- Robot Intelligence* New York, Springer London Dordrecht Heidelberg.2010.
[2]J. Redding, A. Undurit, H-L. Choi, and J. P. How. "An Intelligent Cooperative Control Architecture". *American Control Conf. 2010* . Baltimore, MD, USA. June, 30 -July 2.
[3]A. Richards, J. Bellingham, M. Tillerson, J. P. How. "Coordination and control of multiple UAVs". *AIAA Guidance, Navigation and Control Conf.* Monterey, California. August 2002.
[4]J. Beck, S. Kopp, A. Knoll, and M. Merz. "Functional Architecture for Autonomous Vehicle Operations"*AIAA Infotech@ Aerospace*..Seattle, WA, April 2009.
[5]K. Liang, X. Zhou, K. Zhang, R. Sheng. "Architecture for Autonomous Distributed Real-time Embeded System ". *2008 Int.Symp. on Information Processing*. Moscow, Russia. PP: 568-572. May 2008.
[6]C. Madden, M. Hoen, P. F. Dominey. "A cognitive neuroscience perspective on embodied language for human–robot cooperation". *Brain & Language*. 112 (2010) PP: 180-188.
[7]National Science Foundation. "Information and Intelligent Systems (IIS): Core Programs." NSF 11556. Dec 5, 2011.
[8]Air Force Office of Scientific Research and Arlington VA. "AFOSR Broad Agency Announcement (BAA-AFOSR-2011-01)." Dec 5, 2011.
[9]An-army-of-robots-will-explore-new-worlds .URL : http://en.wikinoticia.com/culture-science/Science/10439. [Cited May 2012]
[10]B. J. Baars, N. M. Gage Cognition, *Brain, and Consciousness: Introduction to Cognitive Neuroscience*. Academic Press. 2007.
[11]J. S. Albus, *Brains, behaviors and robotics*. McGraw-Hill. Peterborough, N.H., USA.1982.
[12]Hippocampus .URL: http://en.wikipedia.org/wiki/Hippocampus. [Cited May 2012]
[13]The Central Nervous System.http://www.kidsbiology.com/human_biology/nervous-system2.php. [Cited May 2012]
[14]K. M. Passino, *Biomimicry for Optimization, Control, and Automation*. London: Springer-Verlag.
[15]Xudan. Xu, Ping. Zhang, Fan. Yang, "The architecture research for a multi-agent team cooperative mission management system" . *Control and Decision Conf. (CCDC)*, 2011, 23-25 May. Mianyang, China.
[16]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, ch.11, pp. 265-292, 2011.
[17]J. A. Starzyk, J. T. Graham, P. Raif, A-H. Tan. "Motivated Learning for the Development of Autonomous Systems" . *Cognitive Systems Research*. 14 (2012), pp. 10-25.
[18]J. A. Starzyk and D. Prasad, “A Computational Model of Machine Consciousness” . *International Journal of Machine Conciousness*, vol. 3, Issue 2, Dec. 2011, pp. 255 – 281.
[19]E. Bicho, W. Erlhagen, L. Louro, E. Costa e Silva. "Neuro-cognitive mechanisms of decision making in joint action: A human–robot interaction study" . *Human Movement Science*. 30 (2011) PP: 846-868.
[20]J. Fisher, D. A. Lawrence and J. Zhu, “Autocommander - A Supervisory Controller for Integrated Guidance and Control for the 2nd Generation Reusable Launch Vehicle,” *AIAA Guidance, Navigation and Control Conf.*, Monterey, CA, Aug. 2002.
[21]Tulving, E., “Episodic and semantic memory”, *Organization of Memory*, edited by E. Tulving, and W. Donaldson, Academic Press, Inc, New York, NY, 1972, pp. 382 – 402.
[22] R. A. Brooks, “Intelligence without representation,” *Artificial Intelligence*, 47, 1991, pp. 139– 159.
[23] R. Pfeifer, and C. Scheier, *Understanding Intelligence*, MIT Press, Cambridge, MA, 1999.
[24]J. A. Starzyk, "Motivation in Embodied Intelligence," in *Frontiers in Robotics, Automation and Control*, I-Tech Education and Publishing, Oct. 2008, pp. 83-110.
[25]J.A. Starzyk, “Mental Saccades in Control of Cognitive Process”, Int. Joint Conf. on Neural Networks, San Jose, CA, July 31 - August 5, 2011.
[26]NeoAxis -- all-purpose, modern 3D graphics engine for 3D simulations, visualizations and games. http://www.neoaxis.com/
[27]"Army - Robots Detonate Roadside Bombs in Iraq." URL: http://www.youtube.com/watch?v=Yx-U-r5DE7E&feature=player_embedded&noredirect=1. [Cited May 2012]
[28]"Robot swarm invades from the ground and air". URL : http://www.youtube.com/watch?v=Fy1lYF53d8Y. [Cited May 2012]