A Cognitive Architecture for Knowledge Exploitation

G.W. Ng, Y.S. Tan, L.N. Teow, K.H. Ng, K.H. Tan, R.Z. Chan

Cognition and Fusion Laboratory, DSO National Laboratories, Singapore
{ngeewah,tyuansin,tloonin,nkhinhua,tkhenghw,cruizhon}@dso.org.sg

Abstract

A cognitive architecture specifies a computational infrastructure that defines the various regions/functions working as a whole to produce human-like intelligence [1]. It also defines the main connectivity and information flow between various regions/functions. These functions and the connectivity between them in turn facilitate and provide implementation specifications for a variety of algorithms. Drawing inspirations from Computational Science, Neuroscience and Psychology, a top-level cognitive architecture which models the information processing in human brain is developed. Three key design principles [2] inspired by the brain – Hierarchical Structure, Distributed Memory and Parallelism – are incorporated into the architecture. A prototype cognitive system is developed and it is able to bring to bear different types of knowledge to solve a problem. It has been applied to object recognition in images. The cognitive system is able to exploit bottom up perceptual information, top down contextual knowledge and visual feedback in a way similar to how human utilizes different knowledge to recognize objects in images.

Introduction

A cognitive architecture specifies a computational infrastructure that defines the various regions/functions working as a whole to produce human-like intelligence. It also defines the main connectivity and information flow between various regions/functions. These functions and the connectivity between them in turn facilitate and provide implementation specifications for a variety of algorithms. There exist a number of excellent cognitive architectures but many have overlooked the importance of biological validity.

Many artificial intelligence (AI) techniques and computational theories have been developed over the last few decades. However, the vast majority of them focus on modeling only specific aspects of human intelligence. Hallmarks of human intelligence, such as robustness and adaptability, are usually "programmed" into systems and not as outcomes. To achieve human-like intelligence, we need to look into the seat of human intelligence – the human brain. We need to understand the different parts of the human brain, how they are connected, what kind of information they process and how they process it. Advances in medical science, especially Neuroscience, over the years have allowed us to answer some of these questions. With the help of more sophisticated measuring devices such as functional Magnetic Resonance Imaging (fMRI), Neuroscience has provided some insights into this area. Although current understanding of the biological aspects of human brain is still quite limited, we can draw inspirations from what can be observed about it. In other words, we can try to model the behaviors of Man, and to a certain extent, the human brain. It is in this aspect that psychology plays a part.

Drawing inspirations from the fields of Computational Science, Neuroscience and Psychology, a top-level cognitive architecture is developed. Various key parts of the human brain and their functions are identified and included in the design. Some of the desired behaviors are set as design principles. The cognitive architecture also models information processing in the human brain. The human brain is able to process information in parallel and is able to bring to bear different types of knowledge, distributed throughout the brain, to solve a problem.

The top-level cognitive architecture design and the design principles will be presented here, together with a description of a prototype cognitive system developed based on this design. This is followed by a discussion on how the cognitive system has been applied to object recognition in images, using contextual knowledge and visual feedback, in a way similar to how a human recognizes objects in images.
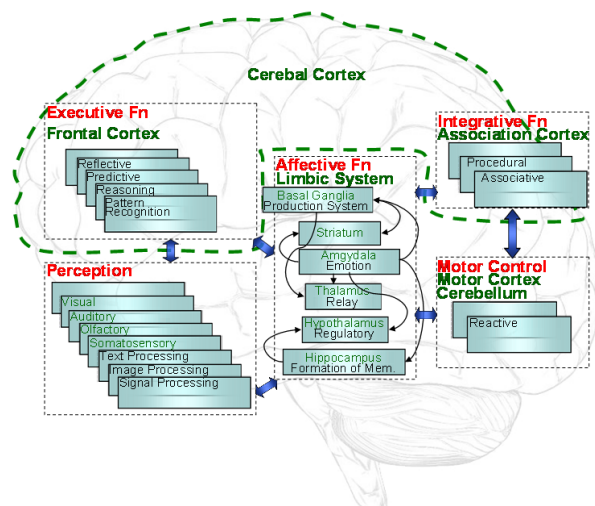
Top-level Cognitive Architecture

Figure 1: Top-level Cognitive Architecture Design

Core Modules

Five core regions in the human brain, namely, Frontal Cortex, Perception, Limbic System, Association Cortex and Motor Cortex, are identified and shown in Figure 1. Each of these five regions represents a class of functions or processes in the brain. The corresponding classes of functions are Executive Functions, Perception, Affective Functions, Integrative Functions and Motor Control, respectively.

Pre-Frontal Cortex (Executive Functions). The prefrontal cortex (PFC) is the anterior part of the frontal lobes of the brain. It has been implicated in planning complex cognitive behaviors, personality expression, and moderating correct social behavior. It is important when "top-down" processing is needed; that is, when behavior is guided by internal states or intentions [3]. The basic activity of this brain region is considered to be orchestration of thoughts and actions in accordance with internal goals. Executive function relates to abilities to differentiate among conflicting thoughts, determine good and bad, better and best, same and different, future consequences of current activities, working towards a defined goal, prediction of outcomes, expectation based on actions, and social "control".

Perception. Perception is the process of acquiring, interpreting, selecting, and organizing sensory information.

Limbic System (Affective Functions). The limbic system [4] is a term for a set of brain structures including the hippocampus and amygdala that support a variety of functions including emotion, behavior and formation of long term memory.

Association Cortex (Integrative Functions). John Hughlings Jackson first proposed in the 1870s that the cortex is organized hierarchically and that some cortical areas serve higher-order integrative functions that are neither purely sensory nor purely motor but associative [5]. These higher-order cortices are what we call today the association areas, associating sensory inputs to motor outputs and performing mental task mediating between sensory inputs and motor outputs. Although the association areas are located at various parts of the brain, we have grouped them together as a functional region.

Motor Cortex (Motor Control). It is a term that describes regions of the cerebral cortex involved in the planning, control, and execution of voluntary motor functions.

Key Design Principles

Three main characteristics, Hierarchical Structure, Distributed Memory and Parallelism, of how the human brain works are identified and these characteristics serve as the key design principles for the cognitive architecture. We believe that modeling the different parts of the human brain and applying these principles will give rise to the robustness, speed, adaptability and other features we have come to associate with human intelligence.

Hierarchical Structure. The neurologist Paul MacLean has proposed that our skull holds not one brain but three [3], each representing a distinct evolutionary stratum that has formed upon the older layer before it, like an archaeological site. He calls it the "triune brain". He refers to these three brains as the neocortex or neo-mammalian brain, the limbic or paleo-mammalian system, and the reptilian brain that includes the brainstem and cerebellum. Each of the three brains is connected by nerves to the other two, but each seems to operate as its own brain system with distinct capacities.

The archipallium or primitive (reptilian) brain, or "Basal Brain", called by MacLean the "R-complex" and which includes the brain stem and the cerebellum, is the oldest brain. It consists of the structures of the brain stem - medulla, pons, cerebellum, mesencephalon, and the oldest basal nuclei - the globus pallidus and the olfactory bulbs. In animals such as reptiles, the brain stem and cerebellum dominate. For this reason it is commonly referred to as the "reptilian brain". It keeps repeating the same behaviors over and over again, never learning from past mistakes. This part of the brain is active, even in deep sleep.

In 1952, MacLean first coined the name "limbic system" for the middle part of the brain. It can also be termed the paleopallium or intermediate (old mammalian) brain. It corresponds to the brain of most mammals, especially the earlier ones. The old mammalian brain residing in the limbic system is concerned with emotions and instincts, feeding, fighting, fleeing, and sexual behavior. To this brain, survival depends on avoidance of pain and repetition of pleasure. Physiologically, it includes the hypothalamus, hippocampus, and amygdala. It has vast interconnections with the neocortex, so that brain functions are neither purely limbic nor purely cortical but a mixture of both. As MacLean understands it, this lowly mammalian brain of the limbic system tends to be the seat of our value judgments, instead of the more advanced neocortex. It decides whether our higher brain has a "good" idea or not, whether it feels true and right.

The Neocortex, alternatively known as the cerebrum, the neopallium, or the superior or rational (neomammalian) brain, comprises almost the whole of the hemispheres (made up of a more recent type of cortex) and some subcortical neuronal groups. It corresponds to the brain of primates and, consequently, the human species. The higher cognitive functions which distinguish Man from the animals are in the cortex. MacLean refers to the cortex as "the mother of invention and father of abstract thought". In Man, the neocortex takes up two thirds of the total brain mass. Although all other animals also have a neocortex, it is usually relatively small, with few or no folds (indicating the surface area, which is a measure of complexity and development).
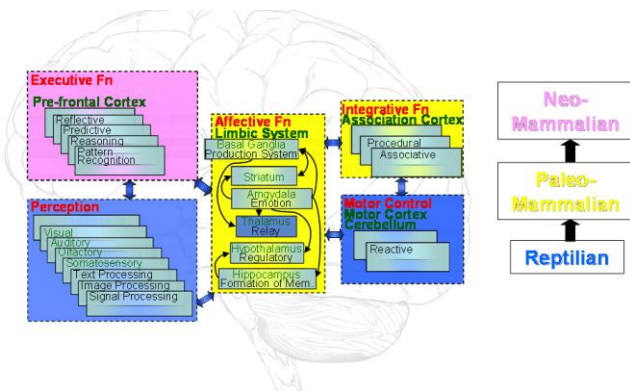
Figure 2: Three levels of Hierarchy in the Human Brain

These three brains form a hierarchy of three brains in one. The cognitive architecture adopts this hierarchical structure in its design to be used as a guide to where various types of knowledge are stored and how information should flow. The various modules in each level of the hierarchy are shown in Figure 2.

Distributed Memory. There are three main types of memory. Semantic Memory consists of facts of the world, disassociated from the place and time when you learned them. Procedural Memory is knowledge about how to do things in the world – it includes your knowledge about how to ride a bicycle, how to type, how to read and understand language, and in general, how to make decisions in selecting actions to achieve goals. Episodic Memory consists of historical episodes or snapshots of specific experiences that are situated in space and time. Studies have shown that memory is not located in any one area in the human brain [6, 7]. Instead, it is distributed throughout the brain. Based on this concept, the cognitive architecture does not have a single module where all the memory or knowledge resides. Each module may have its own memory which it can use to perform its functions or send to other modules when necessary. This will add robustness to the system as it can still function even when some of the functions are down or when knowledge is not complete.

Parallelism. The third key design principle is Parallelism. The human brain does not work in a sequential manner but rather, all the different parts of the brain are constantly running in parallel. This enables the human brain to handle multiple tasks and threads of thoughts at one time. This implies that the brain is able to process different information at the same time. Following this key design principle, the different modules in the cognitive architecture will also be running in parallel. Each module will be developed as an individual running program. The ideal case is to have each of the modules running in one computer in a network. This will allow for true parallelism and hence efficient multi-tasking.

Prototype Cognitive System

A prototype cognitive system (Figure 3) is developed based on the top level design. Some functions from each of the five core regions are developed as modules which form the basic building blocks.

A module is the smallest functional unit of the computational architecture and provides a certain capability. A module is fully encapsulated, with its own knowledge base (distributed long term memory), internal representation schemes and inference methods. Thus a module can be treated like a black box. Other modules in the system do not have to know how it works internally. Each module communicates with other modules either directly or through the Relay (Thalamus) module. Since different modules may have different internal representation schemes, a potential communication problem among the modules may arise in the computational architecture. This problem can be solved by adopting a common representation scheme for all the outputs of the modules.

Modules that perform similar functions are grouped together into classes. For instance, the Perception class comprises of all modules that perform perceptual functions. The reason for grouping similar modules into classes is because different algorithms may be used to find the solution for different problem spaces. By having the concept of classes, each module in the same class can implement just one specific algorithm. This makes the code of each module smaller and easier to maintain. The modules in a class can have complementary, competitive or cooperative relationships. A meta-module for each class may be required to manage the outputs from the different modules within the class.

The prototype system implements each module as an individual executable program. This is in concordance with the parallelism principle of the cognitive architecture.

Description

Perception class: Modules belonging to the Perception class act as receivers to the external world. They take in raw inputs from the external world and process them into useful information. The processed information is then sent to the Relay module for distribution to the rest of the modules in the agent. The current implementation involves a biologically inspired pattern recognition algorithm, Hierarchical Temporal Memory (HTM) [8]. It has an edge over other approaches as it is able to do generalization by exploiting the role of time in vision. In the human eyes, there are short and swift movements called saccades and stops called fixation. We actually make use of these saccades and fixations to visualize and learn the objects we see. This temporal aspect of learning has not been taken into account by many approaches but it is one of the fundamental aspects of HTM that makes it capable of imagery classification.

Motor class: Modules in the Motor class are used to alter both the external environment and the internal state of the

agent. These modules receive instructions from modules such as Selector and apply the necessary actions to the external environment or internal state of the agent.

Association class: Association modules retrieve a list of plausible actions or states when presented with a situation picture. This list of actions or states is associated with the current decision or situation picture. The list is then sent back to the Relay module for further processing by other modules. The current implementation contains a module which builds upon a rule-based engine.

Reasoner class: Reasoner modules analyze situations and proposed actions. They are responsible for higher-level reasoning. The current implementation contains a Dynamic Reasoner module which uses D'Brain [9] for its internal algorithm. D'Brain employs the idea of knowledge fragments and Bayesian reasoning to perform its analysis. The Dynamic Reasoner can be used to fuse different knowledge fragments together.

Selector class: The role of Selector modules is to select an action or a decision from a list of proposed actions or decisions so as to reach the current goals or sub-goals. Currently, the selection process takes into account the probability values provided by the Reasoner modules if they are available. The current implementation contains a FALCON module [10] which enables reinforcement learning in the cognitive system. Reinforcement learning is learning what to do – how to map situations to actions – so as to maximize a numerical reward signal. The learner is not told which actions to take, as in most forms of machine learning, but instead must discover which actions yield the most reward by trying them. Reinforcement learning methods typically have both inductive and deductive aspects: they inductively improve their credibility space on a stage-by stage basis; they deductively select an appropriate response to incoming stimuli using their credibility space. This will enable the Selector module to make better selections over time.

Relay module: The Relay module distributes information to the relevant modules and maintains the current situation picture, in a form of working memory, for all the modules in the system. It functions like the Thalamus in the Limbic System. The current Relay module is able to combine information from different modules and distribute the information to the relevant modules.

Goals Monitoring module: The purpose of the Goals Monitoring module is to produce appropriate sub-goals from the top level goals and then monitor the current situation to check for status of these goals. The status of the goals can be used to update the other modules which may affect their processing of information.

Object Classification in Images

This section will describe how the cognitive system has been applied to object classification in images. Although there has been much research in imagery classification, most algorithms consider each potential target independently and are based solely on measurements of

that target. Due to the nature of the images, the performance of these classification methods generally cannot meet all the operational requirements for accurate classification/recognition.

Human interpreters do not rely solely on the images to do their classification. In reality, they also consider contextual information and inputs from other sources. Hence, regardless of how well the classifier can perform, as long as it does not take into account other information, especially contextual information, users may not have the confidence to use the results of the classification. This is not unlike how we, humans, "classify" objects. We also consider different inputs and contextual information when we are trying to identify objects.
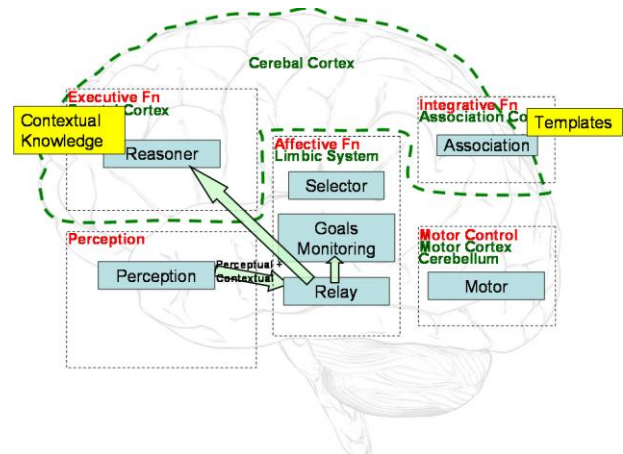

Figure 3: Prototype Cognitive System

Using the Cognitive System

As described previously, the cognitive system is developed based on three key design principles – Parallelism, Hierarchical Structure and Distributed Memory. This leads to certain design features, one of which is the ability to process different kinds of knowledge. This is similar to how Man uses different types of knowledge to solve problems. As mentioned above, there is a need to consider contextual information in the classification process to make it more useful for actual operations. The cognitive system, with its ability to fuse together different types of knowledge, can be used to achieve this.

We, as humans, typically use top-down and bottom-up information to solve problems in a kind of signal-symbol fusion. Contextual knowledge captured in the form of long-term memory is a form of top-down symbolic input while the actual image provides the bottom-up signal information. Contextual knowledge can be seen as a form of prior knowledge which may be learned or gathered through experience. Another top-down information process is feedback to Perception. Previous research has shown that our visual pathways are not unidirectional [11]; in other words, there are also feedback signals to our visual cortex. The system models this by storing templates (the same templates that the Perception module is trained on) in the Association module and retrieving associated templates to send to the Perception module as a visual feedback.

The arrows in Figure 3 show how the perceptual inputs from the image are sent to the different parts of the cognitive system via the Relay module. Certain contextual information may be present in the image itself, for example, a particular formation of objects or other objects in the same image which can help to identify the object of interest. This can be extracted and sent together with the classification results and other contextual information that is outside the image to the other parts of the cognitive system. These form the bottom-up information.
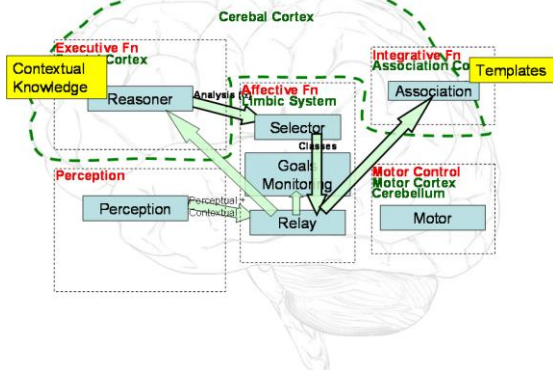

Figure 4: Contextual Information

Contextual knowledge is stored in the Executive Function as shown in Figure 4. The current implementation uses D'Brain as the reasoning engine and the contextual knowledge is captured in the form of Bayesian Network fragments. The HTM output and the contextual information will instantiate some of these fragments which will piece together to form a situation specific Bayesian network. In this way, the bottom-up perceptual inputs are fused with the contextual knowledge. The inference results from the Reasoning engine are then sent to the Selector module. The Selector module will choose the top few classes (classification classes) based on the results and send them to the Association module via the Relay module.

Next, the Association module will retrieve the corresponding templates based on the selected classes. It then sends them to the Perception module, via the Relay module, as feedback to the Perception module. At the Perception module, each template will be "blended" with the original target chip. The blending mechanism is modeled after the human visual recognition process whereby perceived images are adjusted with respect to preconceived templates. Humans model objects and derive the preconceived templates by key features as well as the stabilities of these features. Thus, when we are blending a perceptual input with a potential template, we take into account the features stabilities - features which are more stable are less adjusted. The blended image is then sent to the HTM for classification. This feedback forms part of the top-down information. It is similar to how we retrieve images of objects we have seen before from our long term memory, when we "feel" that the object we are seeing may be one of them. It is important to note here that more than one template is retrieved from the Association module.

When the correct template is used, the feedback should help to boost the confidence that the object belongs to the same class as the template. However, when a template of the wrong class is used, the confidence that the object belongs to this wrong class should be lowered. This can help to prevent the system from being biased to a particular class or self-reinforcing wrongly.
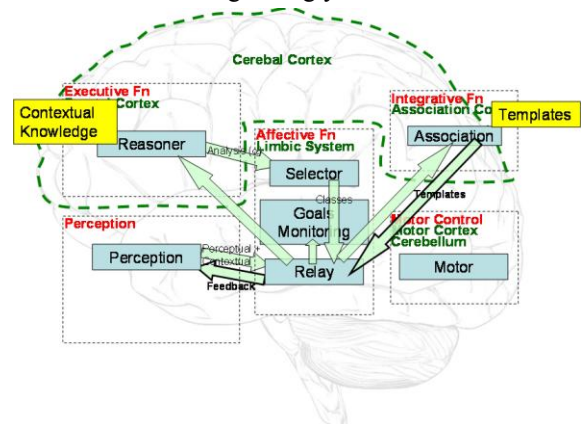

Figure 5: Feedback to the Perception

An Example

An example is used to illustrate how the cognitive system works. The image used for the example is shown in Figure 6. The objective is to identify certain targets in the image. At the start, the user is allowed to enter any contextual information about the image which he may have gathered from other sources. In this example, the user inputs the information that the image is taken near a certain area. The image is passed into the Perception module which carries out a target detection process to find the location of the targets. This is indicated by the square boxes. Next, the Perception module tries to extract whatever contextual information that might exist in the image. The extraction process relies on the contextual knowledge to tell it what to look out for. As formation is one of the contextual knowledge stored, the Perception module tries to see if the targets are deployed in any of known formations. For this case, the targets are found to be deployed in a line formation, as shown by the straight line. Finally, the first target chip on the left is fed into the HTM framework.
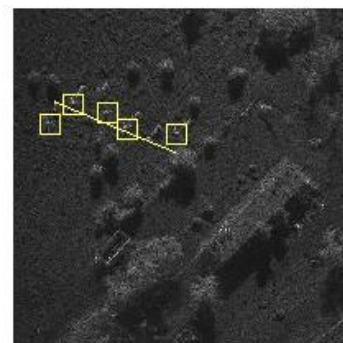

Figure 6: Image

The image processing done to extract the target chip is without the removal of clutter or other image pre-processing. Pre-processing is usually done to "clean up" the image in order to achieve better performance. However, we want to demonstrate how the system can work even when the image quality is far from ideal. As a result, HTM gives a low confidence level of 6% for the correct target class.
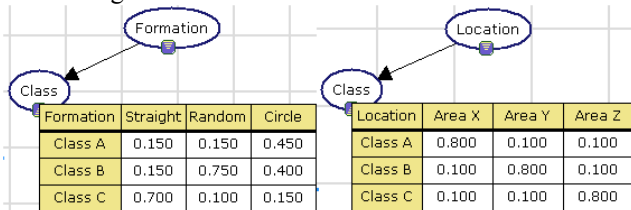

Figure 7: Contextual Knowledge

This result is sent to the Reasoner module which fuses it with the contextual information by instantiating the corresponding contextual knowledge fragments given in Figure 7. As mentioned, two pieces of contextual information were exploited – there is a formation of five targets deployed in a straight line and secondly, this image was taken near a particular area known to contain a certain type of target. One can treat these two pieces of contextual information as belief inputs on the target class from two experts that are conditioned on the information outside the image. The fusion is based on a series of multiplication operations and renormalization [9]. As a result of considering the contextual information, the reasoning engine is able to increase the confidence level for the correct target class to 46%.

This result is then sent to the Selector module which selects the classes of templates to retrieve from the Association module. The selected templates are then sent to the Perception module where it is blended with the original target chip. The blended image is fed into the same HTM framework. Through blending, the template of the correct class is able to fill up some of the missing gaps in the original image as well remove some of the noise. This helps to push the confidence level for the correct target class up to 67%. Finally, the Reasoner fuses this new output from the Perception module with the contextual information to give a high confidence level of 97%. The system stops here as the top-level goal of achieving at least 80% confidence has been met.

Conclusions

A cognitive architecture that models after the human brain information processing is presented here. It identifies core regions of the human brain and functions that exist in each region. Key design principles inspired by the human brain are discussed and used in the cognitive architecture. It is believed that the hallmarks of human intelligence is an outcome of the way the human brain is designed and the cognitive architecture attempts to reproduce these.

A prototype cognitive system has been developed and described here. Various modules from the cognitive architecture are implemented using existing algorithms and programs. One key feature of the cognitive system is its ability to bring to bear different types of knowledge to solve problems and this is demonstrated with an imagery classification example.

Results show that incorporating contextual information and visual feedback in a human-like approach helps to improve the performance of imagery classification. In the example, the confidence of correct classification increases from 6%, when only the target chip is considered, to 97%, when all information are considered.

Like the human brain, the cognitive system is developed to be a generic intelligent system which has many potential applications. It can be used to perform different tasks by feeding the relevant knowledge to the system. Current work includes applying the cognitive system to Computer Generate Forces and Unmanned Ground Vehicle navigation.

References

[1] Newell, A. 1990. *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

[2] Ng, G.W. 2009. *Brain-Mind Machinery*. World Scentific.

[3] Miller E.K., Cohen J.D. 2001. An integrative theory of prefrontal cortext function. *Annu. Rev. Neurosci*. 24: 167-202.

[4] MacLean, P. D. 1990. *The triune brain in evolution: Role in paleocerebral functions*. New York: Plenum Press.

[5] Kandel E.R., Schwartz J.H. & Jessell T.M. 2000, *Principles of Neural Science*. Stamford, Conn: McGraw-Hill.

[6] Lashley, K. S. 1950. In search of the engram. *Society of Experimental Biology*, Symposium 4: 454-482.

[7] Fuster J.M. 1998. Distributed Memory for Both Long and Short. *Neurobiology of Learning and Memory*. Vol. 70, Issues 1-2: 268-274.

[8] Hawkins Jeff. 2004. *On Intelligence*. Times Books, New York.

[9] Ng, G. W., Ng, K. H., Tan, K. H., & Goh, C. H. K. 2006. The Ultimate Challenge of Commander's Decision Aids: The Cognition Based Dynamic Reasoning Machine. In *Proceeding of 25th Army Science Conference*.

[10] Tan, A.H., Carpenter, G.A., and Grossberg, S. 2007. Intelligence through interaction: Towards a unified theory for learning. In *Proceedings ISNN, LNCS4491*, 1098-1107.

[11] Edward M. C. 2004. Feedforward, feedback and inhibitory connections in primate visual cortex. *Neural Networks*, Vol. 17, No. 5-6, pp. 625-632.