

A logical model of consciousness on a neural network system with a simple abstract brain-like structure

Yasuo Kinouchi

*Department of Information systems, Tokyo University of Information Sciences,
Chiba, 265-8501, Japan.*

Abstract

In the world of information technology, many systems with a complex configuration of many computers connected by various networks are currently in operation. A user can easily use one of these complex systems as if it were composed of only one computer. Usually a user does not worry about the complexity of the system configuration. Similarly, we use our tremendously complex brain with ease. Usually we do not worry about the complexity of our brains.

To investigate consciousness and self, a design method based on functional hierarchy and invisible computing, which is effective in complex information system designs^{1,2}, was applied to a design of an autonomous adaptive system that has a similar configuration to a animal brain and uses neural networks as a system element³.

1. Basic configuration and functions

1.1 Basic configuration

The basic configuration of the system was designed to match certain functions of an animal brain, shown below.

1. The system must adapt autonomously to a complex environment without a teacher. Large amounts of information are input to the system from this environment.
2. When adaptation operation begins, the system has basic functions corresponding to the instinctual body-control functions of animals. For example, pleasant and unpleasant rewards can be distinguished.
3. Though neural nodes that operate very slowly compared to normal electric gates are used, the system must decide upon an action as quickly as possible in accordance with the system's current situation and previous experiences.
4. The system must use these experiences as effectively as possible.
5. Only a minimum amount of resources and energy can be used by the system to adapt.

The basic configuration of the system is shown in Figs. 1 and 2. To investigate clearly, the configuration was simplified and abstracted as much as possible.

The system is mainly composed of several functional modules. A concept module (CM) with concept formation functions operates bi-directionally as an encoder of large amount of information expressed by micro-features, and a decoder of recollected group of concepts in the system. A binding module (BM) connects information of an object and its location. A declarative memory module (DM) with an episodic and generalized-rule memory function memorizes the sequence of the system's experiences such as scenes, decided actions, and rewards, and recalls these experiences associatively. In addition, generalized rules are abstracted from these experiences.

An integration and core control module (ICM) connects these modules. The integration part of this module (I-ICM) selects the most important combination of concepts or information in the system at that time on the basis of a mutual vote executed by nodes that represent concepts or emotional states. This function corresponds to global workspace theory (GWT)^{4,5,6}. The core control part of this module (C-ICM) only has an active control function and decides whether to act as a system based on emotional states. Emotion is thought to be information sent automatically from automated part of the system, such as control signals of visceral organs adapting to various situations in case of animals^{7,8}. Function modules above mentioned can be implemented by the method based on Haikonen⁹ and Kinouchi et. al^{10,11,12} using neural network.

1.2 Basic functions in a lower layer

As a main learning method, the system tries to take repeat actions that resulted in a pleasant feeling and to avoid repeating actions that resulted in an unpleasant feeling. The system performs these actions faster and more appropriately through learning. To decide upon an appropriate action as quickly as possible at each time, processing is executed mainly in three paths. Path 1 is the quickest process, which bypasses the BM and ICM. Usually we are not aware of this process. Path 2 is an impulsive process in which an action can be quick, but sometimes inappropriate because declaration memory was not accessed. Path 3 is a reflective process in which an action can be slowly and more appropriately taken because of the use past experiences with associative memory access. The path 1, and 2 roughly correspond to a 2-systems model that explains social behavior as a joint function of reflective and impulsive processes by F. Strack and R. Deutch.¹³

The DM memorizes only information that is selected by mutual vote, decided actions, and rewards in the sequence. At this time, only the short-term memory instantaneously memorizes the content. However, this content is slowly transferred to the long-term memory in sleep mode. In addition, the content of the long-term memory is compressed as a causal rule and stored in the generalized-rule memory.

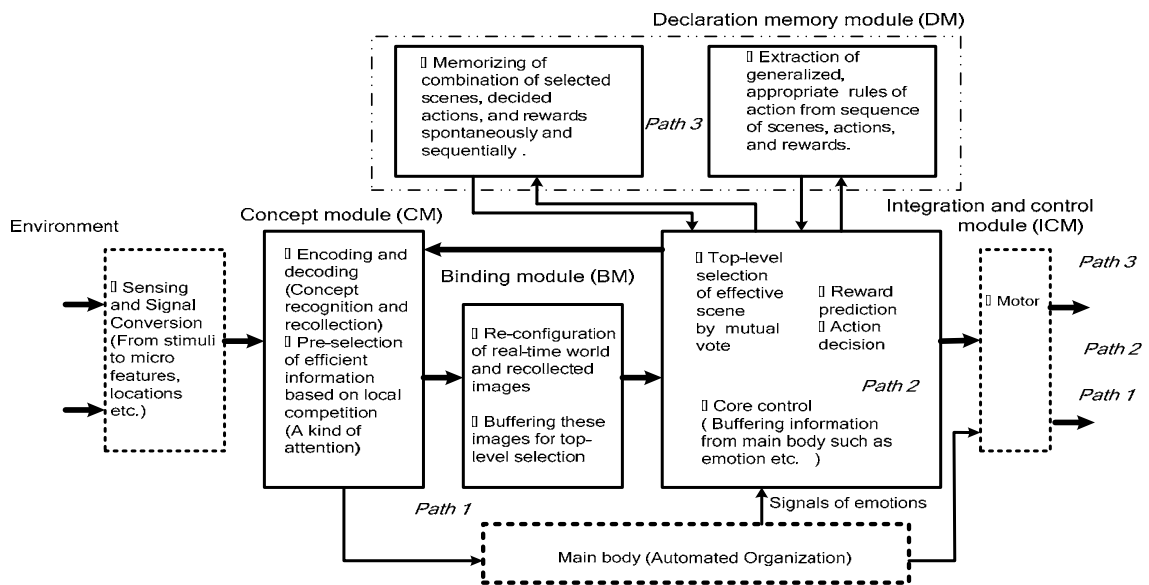


Fig. 1 Basic functions of lower layer

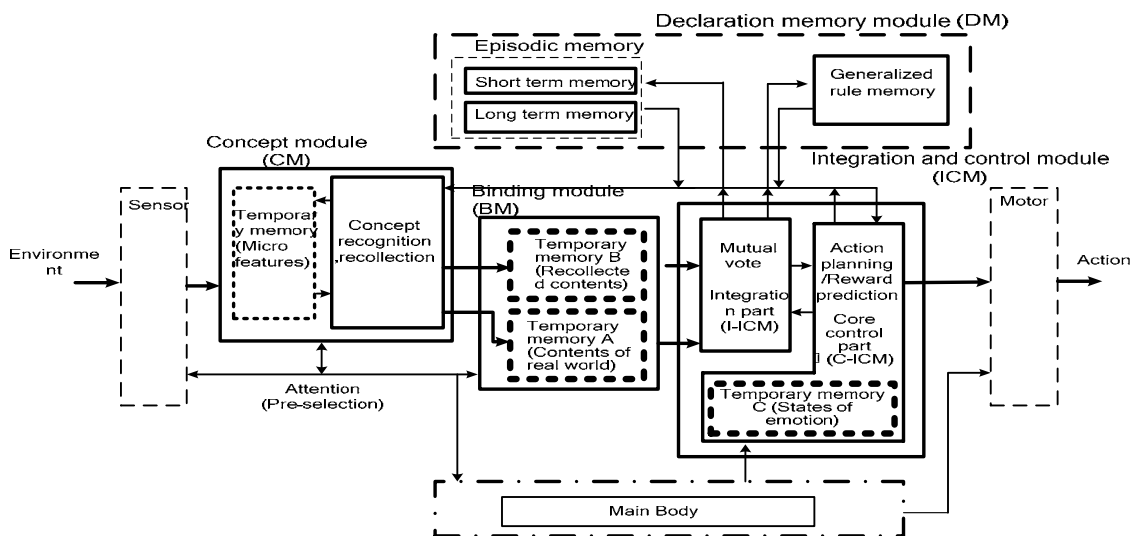


Fig. 2 Configuration of lower layer

More detailed operations of the path 1 and 2 are shown below. First, to decide the action using the entire system, information from the environment and the body is primarily selected by a filtering process corresponding to usual attention. After binding the information of an object and its location, the selected information is stored in temporary memory A. Contrarily, the content recollected previously is decoded by CM, and returned to the information at the micro features level. This information is then primarily selected by attention. Seeing this selected information has no location information because of recollection in the brain, BM stores it in temporary memory B.

In I-ICM, each neural node corresponding information from temporary memories A and B and emotional information from temporary memory C votes mutually. Winner nodes of the vote represent the most important information group for the entire system at that time. Using mutual vote method, flexible selection of important information from a lot of information can be done. An action and reward are predicted based on the information selected here. Then predicted action and reward are checked whether they satisfy conditions that are set by C-ICM according to the emotional states in temporary memory C. When this check is positive, an action is executed. When it is negative, prediction of another action and reward is recollected from DM and compared with the conditions. When an idea, which satisfies the conditions, is recollected, an action is executed. In this case DM is accessed as an associative memory for the result of mutual vote in I-ICM.

2. Self as an upper layer function

For this system to efficiently control and adapt itself, an upper layer was designed over the layer described above, as shown in Fig. 3. The minimum information necessary for the system to control and adapt itself is selected from the lower layer. The information not selected is invisible in the upper layer. The relationship between this selected information is defined as functions of the upper layer.¹⁴ Selection of necessary information is executed according to our daily “awareness”. We daily feel “(1) real-time information of the environment and our body as actual phenomena outside our brain”, and “(2) recollected images and states of mind as mental phenomena inside our brain”. “Self” is defined as a type of function with inputs, output and states. Inputs are real world content in temporary memory A. State are contents of recollection in temporary memory B and information of emotion in temporary memories C. Output of “self” is partially expressed in temporary memory B. “Self” has a blurred region confined by the feeling that (1) is input from the real world to the self, and (2) emerges from inside of itself. The feeling of this emergence inside is formed by the reason that recollected images and emotional content neglect the physical location of the information sources. On the contrary, real substances accompany their physical location information in the real world.

As the upper layer function, “self” recollects past experiences from declarative memory as if it is in the world expressed compactly by decoded information such as scenes (combination of concepts), and decided actions, though it does not directly feel these decided actions. It feels only a kind of images to act, or a change in the body and environment by this action.

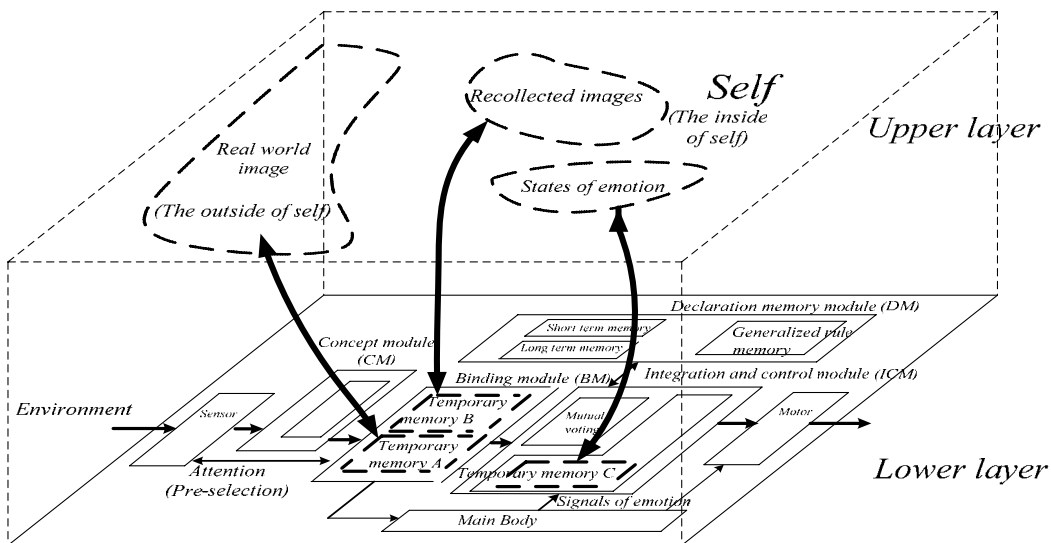


Fig. 3 A function of self in upper layer

3. Discussion

1. Consciousness works effectively when the system flexibly decides on an action by considering the overall situation. In case of path 1, to act as quickly as possible, the system does not use I-ICM, this means that system is unconscious to the action. Then, action of the system becomes a patterned operation without flexibility. On the other hand, to use I-ICM along path 2, it is indispensable that various types of information from a different route are buffered during a certain period. An important matter is not given high priority when the system processes on a first-come-first-serve basis. It is necessary to buffer various types of information, and to select an important matter from among this in a certain period. Additionally, it is necessary to wait for the vote between nodes with the longest distance. That is, a certain fixed time is also indispensable. As for the path 2, a more appropriate action can be done though it is not a high-speed decision. In the path 3, action is more appropriate and rich in variety using past experiences though actions are slowly decided according to the memory access.

2. As “self” cannot be directly aware of its own state with the exception of emotional memory, the contents of “self’s state” has to go around CM and BM to be conscious. This means that “self” can only see itself through; for example, a mirror, and has a property of a “recursive self” based on the basic configuration. We suppose that this function is roughly suitable with the experimental results of Libet.¹⁵ Detailed examinations will be necessary in the future.

4. Conclusion

From the viewpoint of complex information processing system design, a system that autonomously adapts to a complex environment with a brain-like configuration composed of two functional layers: a lower and an upper layer, was designed. Total conscious phenomena are described in detail as adaptive processes with maximum flexibility depending on various functions and information in the lower layer. However, masking certain functions and information in the lower layer based on our daily conscious experiences, “self” is described as a logical function in the upper-layer for simple and efficient adaptation. We are now simulating the configuration on a computer to clarify its operational characteristics in detail. In addition, as the system architecture described here is too simple to be used to construct a brain-like processing system, we are investigating ways to execute higher-level functions.

Acknowledgments

We thank Mr. Shoji Inabayashi of Pacific-Technos Co.Ltd, Mr. Katsumasa Masuda of Tokyo University of Information Sciences, and Dr. Kenji Itoh of the Graduate School of Medicine, University of Tokyo for their useful suggestions.

References

1. Gray, J., and Reuter, A. (1993). *Transaction processing: concepts and techniques*. Morgan Kaufmann
2. Norman, D. (1998). *Invisible Computer*, MIT Press
3. Kinouchi, Y., Masuda, K., and Inabayashi, S.(2008). A model of consciousness and self based on simple abstracted brain-like neural network system. *ASSC 12th Annual Meeting*, June 2008
4. Baars, B. (1988). *A cognitive theory of consciousness*. Cambridge University Press
5. Franklin, S. (2005). Perceptual and Learning: Recognizing, Categorizing, and Relating, *Symposium on developmental robotics* (pp.1-46), AAAI
6. Shanahan, M. (2006), A cognitive architecture that combines internal simulation with a global workspace, *Consciousness and Cognition* Vol. 15 , 433-449
7. LeDoux, J. (2002). *Synaptic Self-How our brains become who we are*, Penguin
8. Damasio, A. (1994). *Descartes' error-Emotion, Reason, and the Human Brain*, Putnam's Sons,
9. Haikonen, P. (2007). *Robot Brains-Circuit and Systems for Conscious machines*, Wiley, West Sussex
10. Kinouchi, Y., Masuda, K., and Inabayashi, S. (2006). An Architecture for an Adaptive System with Primitive Concept Formation and Consciousness. *International Joint Conference on Neural Networks (IJCNN'06)*
11. Kinouchi, Y., Masuda, K., and Inabayashi, S. (2006). Multi-module neural network system with concept formation and primitive consciousness, *Proc. Artificial Life and Robotics (AROB11th)*.

12. Kinouchi, and Y., Sasyo, T., Mizutani, M., and Inabayashi, S. (2001). A structure of an associative memory system with concept formation based on mutual feature reactions, *International Joint Conference on Neural Networks (IJCNN'01)*
13. Strack, F. and Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior, *Personality and Social Psychology Review*, Vol. 8, No. 3, 220-247
14. Kinouchi, Y., and Komiyama, T. (2005). A Study of the basic concept of information in a complex system, *Proc. Artificial Life and Robotics (AROB10th)*
15. Libet, B. (2004). *Mind Time: The Temporal Factor In Consciousness*, Harvard University Press