

# Model of symbolic looking procedure acquisition process in navigation learning task

Takashi Omori

Hokkaido University, Sapporo 060-8828, Japan, omori@complex.eng.hokudai.ac.jp

Akitoshi Ogawa

Hokkaido University, Sapporo 060-8828, Japan, akitoshi@complex.eng.hokudai.ac.jp

## Abstract

In this paper we propose a model of symbolic procedure formation in the brain and its control architecture. In the model, a procedure in the brain is composed of multiple functional parts that realize subfunctions of the required computation. An internal attention vector sequence selects and activates these functional areas, and the processing circuit is realized. To evaluate the plausibility of the model, a computer simulation for acquisition of the reinforcement learning procedure for a navigation task is conducted. Two types of agents with different functional parts have acquired the learning procedure, and the relation of the result with the development of individual is discussed.

## 1. Introduction

Higher cognitive functions of the brain, such as language and planning, are said to be symbolic. Piaget interpreted the developmental process of infants as the development of symbolic processing ability. But when asked “What is a symbol in the brain?”, we cannot provide an answer right now. In the engineering field, symbolic processing is well known computational process. But in the real world, it is not suitably applicable due to, for instance, the symbol grounding problem. On the other hand, it is obvious that brain symbolic processing is grounded to the real world. The understanding of brain symbolic processing is an important research issue.

The development of intelligence in infants gives us a hint towards resolving this problem. Adult intelligence is very complex and is difficult to model. Studying infant intelligence reveals the step-by-step acquisition process of adult intelligence. By tracing the steps and formulating a computational model of each stage, we may reach a stage never reached before without complication.

In this report, we study the mechanism of intelligence that behaves as symbolic, but actually based on a continuous computational process. For its realization, we first study the acceleration phenomenon of infant word acquisition process, and propose a hypothetical model of the mechanism of mental procedure formation for problem solving in the brain. Then, as an example, we show a possible design for the Q-learning procedure acquisition model using a computer simulation of a navigation problem. Here, we discuss on the acquisition of procedure, and not on the

learning. In this research, Q-learning is just an example of procedure.

In the model, the internal procedure is composed of a sequence of subfunctional module combinations. The functional module combination is commonly observed in the brain as the selective activation of task related cortical areas. When we observe the system behavior from the outside, the search process of the combination resembles a symbolic thinking process. Finally, we discuss how our model explains infant intelligence development. In our view, the primary stage of intelligence development can be understood as the accumulation of mental functional parts and the improvement of their combination search skills. We discuss the importance of the mental procedure in addition to the physical body for the realization of grounded symbols.

## 2. Symbolic looking processing in brain

### *Concept of symbolic processing*

Although the symbolic processing mechanism in the brain is not clarified as yet, there are some features that are commonly observed in the so-called symbolic behaviors. The most typical ones are (1) handling of internal representations, (2) discrete handling of discrete representations and (3) inclusion of the search process. On the other hand, the processing of a task that initially requires intentional search becomes automatic after a few iterations. The automatization of the mental process is essential for development to realize the stepwise increment of processing power. A model of symbolic processing development would have to explain the origin of the discreteness, the nature of handling, and the process of automatization.

### *Implication from language development*

Now we consider the case of vocabulary acquisition. In the development of infant language ability, there is a phenomenon of word acquisition acceleration called “vocabulary spurt” that starts around 20 months from birth. As this behavior is not reported in chimpanzees, there should be some mechanism behind it that is specific to humans and human intelligence. As one of the mechanisms, some learning strategies called “bias” have been reported (Imai, 1999). The object whole bias, for example, is the tendency to attach a name to the

entire object, not a part or an attribute such as color or size of the object. It is reported that an infant has some rules for shape bias or category bias that enable quick learning of words. However, these rules do not apply always. The application of each bias is controlled depending on the situation. Infants never apply shape bias to an object that does not have rigid shape, such as clay or liquid. That is, the brain internal behaviors that correspond to various word learning biases are paired with the conditions of application. In real conditions, one of the biases that are suitable for the situation is selected and applied.

From the fact that the infant word learning process is slow at the initial stage of word acquisition, it is obvious that an infant acquires the application situations of the learning rules and internal learning behaviors itself in the process of language learning.

Thus we ask, what is the brain mechanism of the internal behavior that actually represents the learning biases? As the selected bias differs depending on the situation, one simple way is that the brain automatically activates a corresponding neural circuit when it recognizes a specific situation. By increasing the number of situation-learning action pairs through the language learning process, the brain system increases its operable situations and accelerates the word learning rate (Omori and Shimotomai, 2000).

Then, what constitutes the body of internal actions in the brain? And how does it correspond to symbolic processing in brain? In the next section, we propose brain architecture model that realizes the symbolic internal process depending on a situation.

### 3. Procedure representation by functional parts combination

#### *Computational procedure in brain*

From the noninvasive measurement of the brain, activation of different cortical areas depending on a task is reported. As different cortical areas are assumed to have different functions, we can imagine that the brain system combines different functional parts, cortical areas, to form the brain circuit that finds an answer or proper output for the input of the moment.

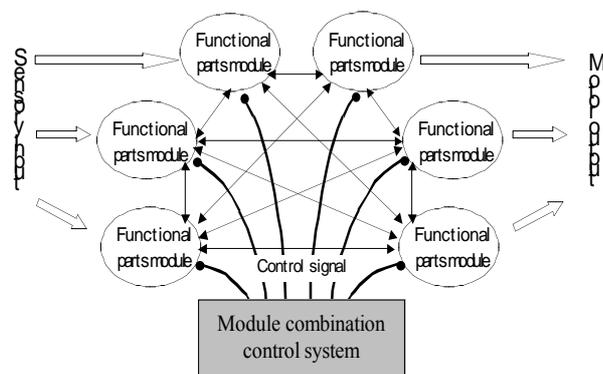


Figure 1 Schematic of computational procedure in brain

For the realization of the combination process, we introduce the notion of internal attention that activates each cortical area separately. Due to the neural connections that exist between most of the cortical areas, the activation of necessary cortical areas automatically and immediately leads to the formation of a neural circuit that processes an input signal and generates an output or internal state without any change of the connections. The internal attention becomes a vector that selects necessary cortical areas, functional parts, to realize the processing.

Based on the concept of computation with a conventional computer, it is natural to assume multiple stages of the combinations that lead to the conception of a sequential program. In that case, the cortical processing circuit that is formed by a sequentially changing attention vector computes data in many steps, and the computational procedure that utilizes the result of the preceding processing is realized. This concept of computation is very similar to the procedures of a conventional computer. Thus, we call the computational circuit that is formed by the sequential selection of cortical areas a “procedure” (Figure 1).

From this conceptual model, it is not difficult to explain how the natures of symbolic systems emerge in the brain. In this model, we assumed the presence of functional parts with the control of activation / deactivation by internal attention. The selection of those parts corresponds to the discreteness of the symbolic process. Each time the system encounters new task, it seeks the suitable functional parts and their combination at the moment to find the required procedure. Though we do not mention the search mechanism here, it is natural to assume the process corresponds to the search in a symbolic system. The procedure formation process becomes automatic when the activation vector for the functional parts is found and memorized corresponding to the task situation.

#### *Component of brain procedure system*

The proposed model is composed of three major parts.

- A set of functional parts  $f_i(x)$  that are mutually connected and can be activated/deactivated by the sequence of internal attention vector  $a_i^K(t) = \{0,1\}$ . Here,  $i$  is the index of the parts and  $t$  is the sequence. Between the functional parts, the connection  $w_{ij}$  interconnects the output of  $f_j(x)$  to the input  $f_i(x)$ .
- An attention generator holds a table of a learned attention vector sequence  $A^K = \{a_i^1, a_i^2, a_i^3, \dots\}$  that corresponds to task situation  $K$ .
- A situation detector that recognizes external input  $E_t$  and internal status  $S_t$  as an already learned situation  $K$  or a new situation.

#### *Behavior of the procedure generation system*

The behavior of the system then becomes as follows.

1. In the initial condition, any internal status  $S_t$  is not activated.

2. Given an external input  $E_I$ , the situation detector recognizes combined input  $[S_I, E_I]$  as an already known situation  $K$ . If no corresponding known situation exists the situation is classified as new.
3. The internal attention vector sequence  $A^K$  is activated and applied to the set of functional parts. The internal attention  $a_i^K(t) = \{0, 1\}$  selects and activates the adopted parts, and realizes the complex function  $S_{II} = f_i \left( \sum_j w_{ij} a_j^K(t) f_j(\cdot) \right)$ . In the case of Figure 2 (a), the realized function can be described as  $S_{II} = f_i(w_{ij} f_j(w_{jk} f_k(E_I)))$ .
4. After the process of the configured processing is finished, the time step of the attention vector  $a_i^K(t) = \{0, 1\}$  proceeds stepwise. In the case of Figure 2 (b), the realized function is described as  $Output = f_n(w_{nm} f_m(w_{mi} S_{II}))$ .
5. Step 4 is iterated until the end of the attention vector sequence  $A^K$ . In some cases,  $a_i^K(t)$  can be single step to realize simple processing.
6. Steps 2 to 5 are repeated. The internal state  $S_I$  works as a context to modify the situation recognition so that the context-dependent procedure is realized. In a sense,  $S_I$  plays the role of a variable in the procedure of conventional computer. It is self-evident that the acquired procedure is domain specific.
7. If the input  $[S_I, E_I]$  is classified as new in the step 2, the attention generator begins to seek a suitable attention vector and its sequence.

#### Procedure search method

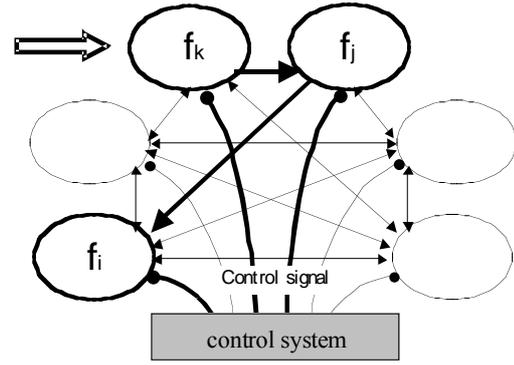
At this point we ask what is the method used to search for the combination of suitable functional parts? The physiological mechanism of module activation in the brain is not yet known. Functionally, a gating of a neural module or a connection between neural modules by an internal selection signal would be sufficient to realize the intended function. The selection of functional parts and the connection between them has the same effect.

The internal attention vector should have a large dimension as there are many modules to be controlled in the brain. Computationally, a combination of associative memory and sequential attention can realize state automata (Omori et al, 1999). Physiologically, there is a hypothesis that the phase relationship between two areas controls the functional relation between them (Bressler and Kelso, 2001).

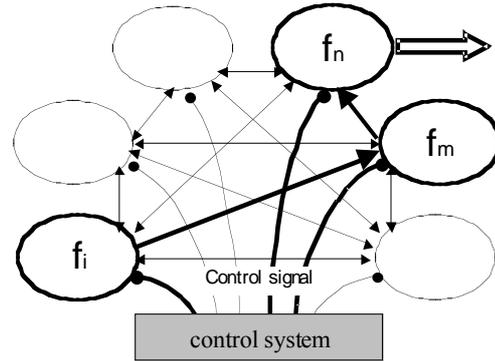
#### Change of procedure in the course of development

The procedure formation ability of a system changes due to the following constraints.

- Number and variation of available functional parts. If the system has sufficient types of functional parts that are necessary for the task of the moment, it can compose the required procedure and solve the



(a) At the first step of the functional vector, input and parts  $f_i, f_j, f_k$  are activated.



(b) At the second step, parts  $f_i, f_m, f_n$  and output are activated.

Figure 2 Example of procedure formation using the set of functional parts and the attention vector generator.

problem. But if one of the necessary functional parts, such as the working memory as described in the next section, is lacking, the system cannot acquire the required procedure and cannot solve the task even if it has many unnecessary functional parts.

- Efficiency of the method to combine the functional parts into a single procedure. Intuitively, humans increase the skill by experience. However the details of the mechanism are not yet known.
- Available length of attention vector sequence. If the system can use a sufficiently long attention vector sequence, it may generate a complex procedure by iterative use of the simple parts.

Our model predicts that the relaxation of these constraints is the mechanism underlying infant intelligence development. The infant brain increases the number and varieties of available functional modules by experience and its intrinsic learning ability. The capacity of a working memory and related memory increases with age, although we do not know detail of the mechanism.

The next question is, how do these constraints work in the procedure acquisition process in the case of real problem solving? To evaluate this issue, we try acquisition of the reinforcement learning procedure using our model.

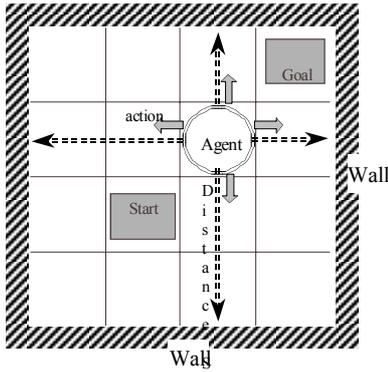


Figure 3. Simulated world for the learning procedure acquisition.

#### 4. Navigation Problem Solving Procedure

##### Navigation and symbolic processing

In the engineering field, path finding is a typical example of symbolic processing that requires a tree search for possible action combinations to acquire a sequence of actions to reach the goal. The search process is composed of the choice of a discrete action, prediction of the result of the action, evaluation of the result, memorization of the action and the evaluation, and reorganization of the best path.

From the viewpoint of our model, the symbolic path search procedure is just same as our combination search behavior of functional parts except that each part is described and implemented as a sequence of computer program. If we can realize the same function by searching for a suitable attention vector sequence, it would be proof that symbolic processing in the brain could be realized by our model. The experiment will be more interesting if we simulate the developmental process by changing the available functional parts.

##### Functional parts for Navigation

The navigation task has been studied extensively through behavioral and physiological study. Although there have been many enlightening findings such as the place cell in the hippocampus, there remain many unclarified behavior of the brain system.

An internal map is necessary for the prediction of a

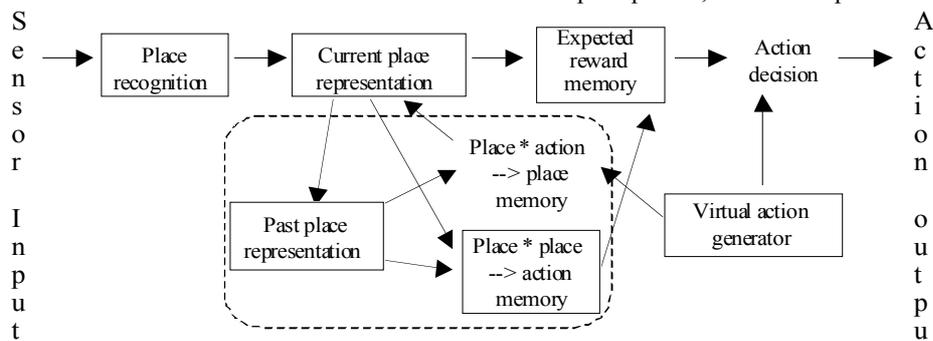


Figure 4 Functional parts of the basic agent and the developed agent. The parts in the dashed circle are not used by the basic agent.

traveling action result. The function of a map is realized by a combination of (1) a recognition of current position from sensory input, (2) a representation of current position independent of sensory input, (3) a prediction of next position by virtual action, (4) a temporal memory of former position, and (5) a memory of the action required to travel from one position to the next position. If we properly combine these functions, we can realize the planning function using the environmental map. Here, the function (1) is realized by a recognition mechanism, (2) and (4) are realized by a working memory, and (3) and (5) by a memory of the sequence. The function of navigation that appears symbolic from the macroscopic view point is realized by the proper combination of functional parts that are not related to the symbolic processing. Actually, the combination process of functional parts would be a computational procedure that is easily implemented in the brain.

##### Procedure Learning Method

Here, we show that the reinforcement learning procedure for the navigation task emerges by the combination of functional parts. The system that can acquire this type of problem solving procedure is expected to have the ability to acquire other procedures even for different tasks and in different environments.

Figure 3 shows the simulated world in which the agent learns goal-reaching behavior. The 4 by 4 grid is enclosed by a wall. The agent moves step by step in four directions and a reward is given when the agent reaches the goal.

In the experiment, we prepare two types of agents. One is the basic agent that has functional parts that are minimum to realize Q-learning. Another is the developed agent that has functional parts that are necessary to realize an environmental map in addition to those of the basic agent (Figure 4). The parts of the developed agent are designed so that the agent can realize prediction-based learning if those parts are used in proper combination with proper timing.

For the attention generator, we used a two-step sequence attention vector. Each bit of the vector controls the on-off state, in use or not in use, of connections between functional modules. By the use of a two-step sequence, we can expect the outcome of a

procedure that uses internal states one time between input and output. We use a genetic algorithm (GA) as the method of searching for the attention vector. The vector data  $a_m^K = \{0,1\}$  that designates the on-off state of the connections is used as the gene. Given a gene, the internal procedure is defined and the fitness of each individual can be measured by the inverse of the learning time of the agent. In this study, we defined the learning time as the number of action steps from the initial state until the agent has taken the shortest goal-reaching path three times consecutively (Ogawa and Omori, 2001).

#### Functional parts for the basic agent

The basic agent has functional parts of sensory input recognition, current place representation, Q-value memory, virtual action generator, and action decision system (Figure 4). Figure 5 shows neural network implementation of those parts and their interconnection. In this implementation, we used the gating of the interconnection instead of activation of the functional module.

The sensory input recognition part recognizes input signal  $E_i$ , and the recognition result  $CP_j$  corresponds one-to-one with the location in the simulated world. The current place representation is simply a set of memory cells  $CP_j$  without any relation between the locations. In our simulation, we prepared 16 cells and their connection for input recognition corresponding to their simulated world location. The recognition of the place from the sensory input is easy to be learned by competitive learning. We did not assume a sensory aliasing situation.

The connection  $V_{ij}$  between the  $CP_j$  and an action representing cell  $A_i$  memorizes the expected reward when the agent takes an transfer action  $i$  at place  $j$ . When it is used for the action choice in the goal reaching task, an association from  $CP_j$  to  $A_i$  and competition between the action cells select the

maximum Q-valued action at the place.

$$\tau \frac{d A_j}{dt} = -\sum_{k \neq j} A_k + \alpha \sum_j V_{ij} CP_j$$

The Q-learning mechanism is embedded by hand in the expected reward memory. Conventional Q-learning equation is always applied to  $V_{ij}$  when the agent takes action, and the Q-value evaluation is given. In the initial state of learning, all  $V_{ij}$  values are zero.

In the action decision layer, the actual action coding cell  $AD_i$  receives input from the  $A_i$  and a random value generator cell  $R_i$  through the attention-gated one-to-one connection. The lateral competition within the  $AD$  cells selects the most activated action based on the effect of the Q-value-based action  $A_i$  and the randomness.

$$\tau \frac{d AD_i}{dt} = -\sum_{k \neq i} AD_k + a_m^K A_i + a_n^K R_i$$

In the attention vector search by GA, half the agents with higher fitness are used for the next generation production. Crossover pairs are decided in the order of 1st and 2nd, 2nd and 3rd ... and so on based on their fitness value. One-point crossover is used, and crossover point is chosen at random. The mutation rate is 0.05. Each individual is set to the initial state, and its learning time is evaluated. After some generations of attention vector search, the agent acquired a typical Q-learning procedure as was expected. As the basic agent does not have an environmental map nor working memory parts, its action was determined depending on immediate sensory input, that is, selecting an action based on the looking up table of the Q-value.

#### Functional parts for the developed agent

In addition to the functional parts of a basic agent, the developed agent has functional parts of a working memory of one step past self location  $PP$ , a place-action to place memory  $PAP$  and a place-place to action memory  $PPA$  that can learn and represent the map of the environment (Figure 4, Figure 5). The map-learning behavior of these memory parts are embedded by the researcher. They are updated after each action by the following equations.

$$PP_i(t) = CP_i(t-1)$$

$$\Delta PAP_{ijk} = PP_i \times A_j (PAP_{ijk} - CP_k)$$

$$\Delta PPA_{ijk} = PP_i \times CP_j (PPA_{ijk} - A_k)$$

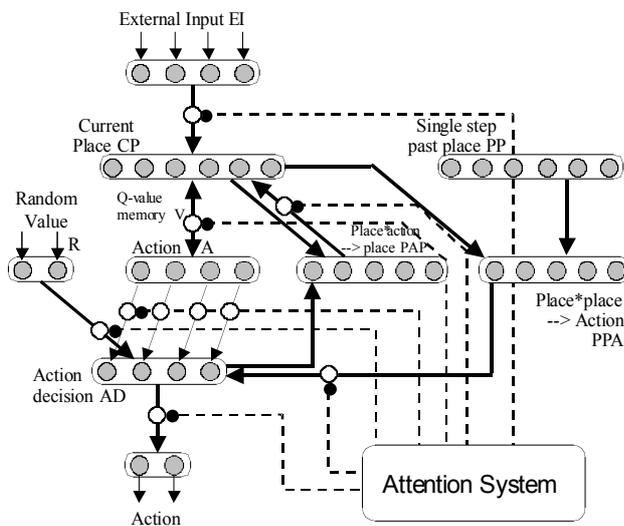


Figure 5 Neural network representation of developed agent

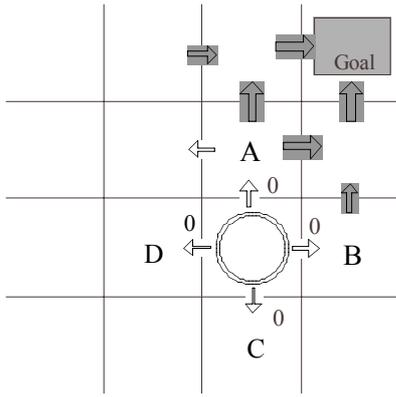


Figure 6 Procedure of prediction-based learning by the developed agent.

Figure 6 shows the learning procedure acquired by the developed agent via the GA search. As the agent has the working memory and environmental map, it acquired the learning procedure that makes use of a single-step-forward Q-value without knowing the Q-value of the current position. This accelerates Q-learning at the border of an already learned area and an unlearned area.

Figure 7 shows the change of maximum fitness value within individuals over generations. As the attention vector is short, eight bits for the basic agent, the optimal procedure for the basic agent is found at the first generation and the fitness value does not change with the generation. Conversely, the vector length of the developed agent is 16, it required five generations to find the optimal procedure of the developed agent, which makes use of a single-step prediction and does not use the random action generator. As the GA algorithm is exactly the same in both agents, the reason for the performance difference should due to the difference of functional parts.

## 5. Discussion

In the experiment, the two types of agents have different functional parts, which induced imbalance of controllable parameters. However, if we add unnecessary parts for this task to the basic agent, and counterbalance the number of parameters, we will get the same result as long as the added parts are not useful in any of possible parts combination. Of course the added parts may increase the possible range of realizable procedures in different tasks and different environments. Thus, the number and variation of functional parts are important for the adaptability of the agent, but not in this task.

In this study, we assumed that the developed agent acquired the additional functional parts somewhere along its developmental course. Then, how are the functional parts acquired along the developmental course? In the real brain, the functional parts are assumed to be self-organized through interaction between the external input and the intrinsic neural learning mechanism. What we have shown in this report

is the possibility of behavioral change by changing the functional parts without any change in the attention control system. This is a possible aspect of development in humans and artificial agents. Although there remains a possibility that the change of the control system occurs simultaneously with development, it is natural to think that most of the infant developmental change is caused by the increase of available functional parts.

If so, what causes the change of functional parts simultaneously with development and evolution? The theoretical study of unsupervised learning says that the characteristic of self-organization changes with various parameters of the learning system, such as the number of cells, range of lateral inhibition, latency of excitation and so forth. The learning dynamics of self-organization that govern the temporal scale of change is also important to explain the gradual change in infants. It is likely that the emergence of a processing module in the infant brain is dependent on these innate neural system parameter settings and on postnatal sensory experiences that also control learning. We have to consider the interaction between the self-organizing innate system and postnatal environment to understand the change of infant behavior.

Our model does not include the formation process of each functional part. As the parts are assumed to correspond to cortical areas in brain, the parts formation model will have some relation to the theory of cortical functional area formation. But little is known about a model of the function (Omori, 1996).

In engineering viewpoint, we are proposing a partial model of grounded symbol emergence (Hanad, 1990). The claim that an interaction with the physical world by a body causes resolution of the symbol grounding problem is correct. But from our model, we consider that both the body and processing system are necessary for the resolution. The symbolic process in our model cannot emerge from a conventional AI model and physical body combination.

## 6. Conclusion

In this report, we explained the model of symbolic procedure formation in the brain by functional parts combination, and showed its effectiveness for the navigation learning procedure acquisition. The computer simulation result has shown that the agent can acquire a classical reinforcement learning procedure and its variation if it is given sufficient functional parts.

With the model, the following phenomena in symbol processing can be explained. (1) The search process of a new task diminishes after a suitable procedure for the situation is found. After the parts combination is memorized, the procedure begins to be activated automatically in the situation. (2) A rather complex computation procedure that uses intermediate results and iteration can be realized with a neural circuit, if the condition of the iteration is correctly encoded in the situation detector. On the other hand, we have not found a model that explains the formation of each functional part.

From the viewpoint of our model, intelligence evolution is the change of innate parameters in the brain, and it decides the range of achievable functions of the brain. Conversely, intelligence development is the process of functional parts acquisition and their combinations for each recognizable situation (Figure 8). There should be many unknown constraints that decide the order of functional parts acquisition. When the number of functional parts exceeds some threshold, adaptability of the individual to its environment explodes, and we feel the individual is intelligent. This is our current image of intelligence. The next problem involves model construction for the acquisition of each functional parts.

## References

- Bressler S., Kelso S. (2001) Cortical coordination dynamics and cognition, Trends in Cognitive Science, Vol.5, No.1, p.26-36
- Hanard S. (1990) The symbol grounding problem, Physica D42, p.335
- Imai, M. & Haryu, E. (1999). Word learning without aid from syntax: How do Japanese children learn proper nouns and common nouns? To appear in the Proceedings of the 23rd Boston University Conference on Language Development. Cassandra Press. 277-288
- Ogawa A., Omori T. (2001) The acquisition of space search procedure depending on agent structure, Proc. of ISHF (in press)
- Omori T. (1996) Model of Cortico-Hippocampal Memory Consolidation Process by Accumulation of Case Memories and Reorganization into Maps", Neuro-Science Meeting, Wash.DC.
- Omori T., Mochizuki A., Mizutani K., Nishizaki M. (1999) Emergence of Symbolic Behavior from Brain Like Memory with Dynamic Attention, Neural Networks, Vol.12, No.7-8, p.1157-1172
- Omori T., Shimotomai T. (2000) Simulation of Human Language Acquisition Process by Brain Like Memory System, Proceedings of ICONIP-2000, 428-433, Nov.14-18, Korea

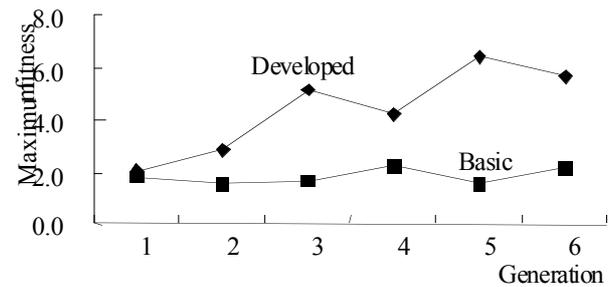


Figure 7 Change of maximum fitness over generations.

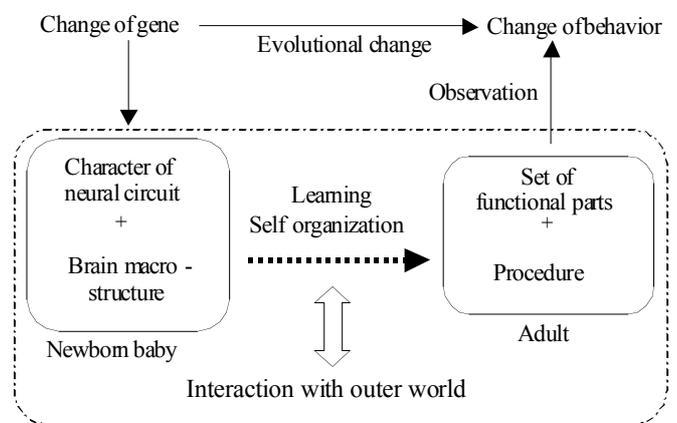


Figure 8 Relationship of evolution and development