A Constructive Approach for Investigating the Emergence of the Primitive Functions of Mind

Kenichi Minoya

Contents

1 Preface 1
1.1 Background 1
1.2 The aim of our study 3
1.2.1 Evolution of the planning abilities...3
1.2.2 Emergence of a Theory of Mind 4
1.3 Description of the Thesis...5
2 Evolution of the Planning Abilities..7
2.1 Introduction...7
2.2 The Model..9
2.2.1 Planner - Beam Search...9
2.2.2 Task - Blocks World Problem 11
2.2.3 Architecture..13
2.2.4 Fitness 14
2.3 Basic Experiments..15
2.3.1 Experimental Setup 15
2.3.2 Results..16
2.4 Evolutionary Experiments 21
2.4.1 Experimental Setup 21
2.4.2 Results..22
2.5 Conclusion and Discussion...27
3 Emergence of a Theory of Mind..29
3.1 Introduction...29
3.2 Three Hypotheses for Investigating the Emergence of Theory of Mind 30
3.3 Functional Model of the Brain 32
3.4 Task and Intention Estimation Model...35
3.4.1 Hunter Task..35
3.4.2 Intention Estimation Model...36
3.5 Experiment 38
3.5.1 Module Network and Function of Each Module 38

3.5.2 Experimental Setup 40
3.5.3 Results 41
3.6 Additional Experiment 44
3.6.1 Module Network and Function of Each Module 44
3.6.2 Experimental Setup 46
3.6.3 Results 46
3.7 Conclusion and Discussion 49
4 Conclusion 53
Acknowledgement 55
References 57
List of Publications 65

Figures

1.1 Scientific approaches to the mind 2
2.1 An example of the state space. Boxes with alphabets represent distinct states. *S* represents the start state. The numbers in boxes represent *h* ...10
2.2 A part of the state space of blocks world problem (*T*=3). "*h*" represents the heuristic value corresponding to the each configuration. 12
2.3 Architecture of an agent (top: basic model; bottom left: nonshared model; bottom right: shared model)..13
2.4 Average action steps to solve a problem in various settings of the table size (*T*), the depth of the problem (*D*), and the planning limit in the basic model 16
2.5 Action steps averaged over 100 trials of the task evaluation when varying the planning limits of two agents where *F*=18 and *T*=3 (top: nonshared model; bottom: shared model) 18
2.6 Average action steps of 100 trials of the task evaluation when two agents took the same planning limit where *F*=18 and *T*=3........................19
2.7 Fitness landscape of the agent B where action steps were averaged over 100 trials of the task evaluation when varying the planning limits of two agents where *F*=18, *T*=3, and w_p=0.1 (top: nonshared model; bottom: shared model) 20
2.8 Difference in transition of the fitness, action steps, and planning limits between the basic model, the nonshared model, and the shared model when the difficulty of the given problem was changed (w_p=0.1)................24
2.9 Average planning limits between 200 to 300 generation in the basic model, the nonshared model, and the shared model when the difficulty of the given problem was changed...25
2.10 Transition of the distribution of the planning limit in the population on a certain trial..26
3.1 Functional model of the brain...33
3.2 Algorithm of the searching system 34
3.3 Hunter task..36

3.4 Higher-order estimation of other's intention in terms of intentional stance developed by Dennett (1987): Agents at level 0 decide their behavior without considering behavior of others; Agents at level 1 assume others as level 0 and decide their optimal behavior; Agents at level 2 assume others as level 1 and decide their optimal behavior 37
3.5 The module network used for the experiments 38
3.6 The transition of the *Solution*$_p$ (black lines), *reward* (pink line), E_p (blue line) and T_p (red line)...42
3.7 The emerged partial networks for processing ToM..............................43
3.8 The transition of the average rewards of *N* agents with and without the attention system, averaged over 10 trials...43
3.9 Active strategy in which agents assume others as level 1 and decide their optimal behavior which informs others the own intention explicitly....44
3.10 The module network used for the experiments 45
3.11 The transition of the activation signals, the link-deactivation signals, and the reward during the learning (7 rounds) 47
3.12 The emerged partial networks for processing ToM............................48
3.13 Conceivable definition of the recursion levels of ToM by using the module network adopted in the experiments...51
3.14 Cooperation among three agents..52

Tables

2.1 An example trace of the beam search algorithm on the state space in Fig. 2.1 when the B = 3. Each superscript represents the value of corresponding h. Each row shows the trace of the search when the planning limit of the agents is different. For example, agents with a planning limit of 2 choose D^2 as a sub goal. Also, agents with a planning limit of 3 choose D^2 or F^2 as a sub goal..11
2.2 Experimental setup of the basic experiment.......................................15
2.3 Experimental setup of the evolutionary experiment.............................22
2.4 Conceptual diagram of the evolution of the planning abilities given by the results of the evolutionary experiments...27
3.1 Experimental setup 40
3.2 Experimental setup 46

1 Preface

1.1 Background

One of the fascinating characteristics of human mind is the ability to imagine one's own future states and the mental states of others. Both future self and present (or future) other cannot be experienced directly and are accessible only through an act of imaginations. These kinds of representations are called *detached representations* which differ from perceptions which are *cued representations*, representations about something that is present here and now (Gärdenfors, 1995). They provide several advantages: if the organism carries a small-scale model of external reality and of its own possible actions within its head, it is able to play trial and error in its internal representation, without risking its life in the real world (Craik, 1943). It comes near to stating the obvious that our ability to understand the mental states of others is also an important when it comes to deal with fellows in social circumstances. This thesis is intended as an investigation of the emergence of the primitive functions of mind by focusing on (1) *planning abilities*, i.e., envisioning various actions and their consequences for getting from start to goal (Gulz, 1991) and (2) *Theory of Mind* (ToM), i.e., understanding of others as having intentional states such as beliefs and desires (Premack & Woodruff, 1978).

For a long time the issue of the nature and role of representation is one of the fundamental questions in philosophy. Thanks to the recent progress in psychology and neuroscience, researchers have attempted to approach this problem in a scientific manner. Recent functional neuroimaging experiments and studies of neurological and psychiatric patients have been producing lots of significant results for understanding our ability to maintain and manipulate mental representations. For example, the left and right parietal cortex has been considered to play an important role for imaging one’s own and another’s mental states, respectively (Allison et al., 2000; Ogawa et al., 2006; Wolpert et al., 1998). It has also been suggested parietal regions predominantly associated with storage functions, while prefrontal areas are related to control processes, such as active maintenance (Levy & Goldman-Rakic, 2000; Oliveri et al., 2001; Shallice, 1988).

These studies are based on the analytic or reductionist approaches that attempt to break down complex systems into a specific part of the system (Fig. 1.1). There is no longer any doubt about the fact that these approaches have been producing lots of valuable results. However, researchers are increasingly aware of their drawbacks. One important drawback in this approach is that the difficulty of analyzing a system grows exponentially as the complexity of the system increases (Evans, 2003). Several functions of mind emerge out of internal recursive interactions between components in the brain and brain-body-environment interactions. Since there are lots of parameters in its system, it is difficult to control them appropriately by the instruction just as scientists would wish (Miwa, 1999). Valentino Braitenberg (1984) argued that if we wish to discover how some system works, it is often easier to do so by building successively more complex models, rather than by attempting to infer the mechanism from mere observation. In general, it might be easier to investigate the mechanisms of the target by synthetic methods than by analytic ones when it comes to complex systems. This is what Braitenberg (1984) refers to as the "law of uphill analysis and downhill invention," that the deductive reasoning of the inventor creates a much simpler (and correct) system than the inductive reasoning of the outside observer. That is to say, in induction one has to search for the way, whereas in deduction on follows a straightforward path (Fig. 1.1).

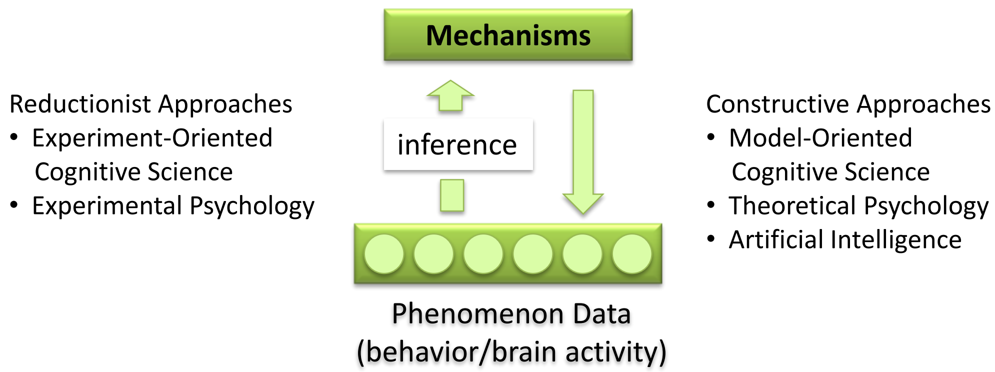


Figure 1.1: Scientific approaches to the mind.

The synthetic or constructive methods are the engineering approaches to the mechanism concerning information processing in the brain. They have been employed to construct models grounded on physiological evidence and some hypotheses, and to analyze the characteristics of the models (Arita et al., 1995). These models not only give informative suggestions to study on the brains, but also are applicable to development of new systems which have facilities for more intelligent and flexible information processing. Furthermore, a new science called *Artificial Life* (ALife) has been established recently which is considered to be an extension of the synthetic approaches. It is developing into a new type of discipline, based on computational construction as its main tool for exploring and producing a science of not only "life as it is" but also "life as it could be" (Langton, 1987). Accordingly, it allows scientists to gain a more profound understanding about the fundamental aspects of life and intelligence than would be possible if their research focus was limited to natural systems only (Bisig and Pfeifer, 2008).

1.2 The aim of our study

This thesis investigates the emergence of the primitive functions of mind, focusing on the planning for future events and general mind reading ability, so called Theory of Mind (ToM), by using a computational model based on the ALife approach. The first goal is to define the conditions for the emergence of planning abilities. The second goal is to investigate the mechanism of the emergence of the ToM by modeling the brain at the functional level.

1.2.1 Evolution of the planning abilities

Human beings have behavioral flexibility based on a general faculty of planning for future events. It has been said that the prefrontal cortex, known to be critically involved in planning abilities, has been especially enlarged through the human evolution than other brain areas (Deacon, 1997). Large brains are extremely costly both to maintain and evolve. Therefore, in a niche where there is little to use planning abilities, it might have a relatively small impact on evolution of it. In chapter 2 we consider the question: what kinds of environment contribute to emergence of the planning abilities?

Natural selection has been considered as one of the most widely held mechanisms to explain the emergence of living creatures' complex characteristics. Evolutionary psychology has been attempted to explain psychological traits such as emotion, cognition, and planning as adaptations as the functional products of natural selection or sexual selection. Recent studies have indicated that ecological problems, such as getting a resource from the environment and avoiding capture by a predator, drove the evolution of intelligence of human (Byrne, 1997; Darwin, 1871; Hill, 1982; Potts, 1998; Tooby & DeVore, 1987). A different approach to the problem of the evolution of intelligence of human involves the consideration of the social aspect, such as the necessity of dealing continually with our fellow humans (Alexander, 1971, 1990; Dunbar, 1998; Humphrey, 1976).

Considering all of the above factors, this chapter explores the dynamics inherent in the mechanism of the evolutionary acquisition of the planning abilities, focusing on the benefits of the planning and the costs of it. The first goal of this chapter is to investigate the properties of the problems which drove the evolution of planning abilities. The second goal is to explore how the difference in a way of interacting with other individuals affects an evolution of the planning abilities.

1.2.2 Emergence of a Theory of Mind

In the third chapter, we discuss a more social aspect of human cognition: Theory of Mind (ToM). Its origins can be traced back in extant non-human primates; ToM probably has emerged as an adaptive response to increasingly complex primate social interaction (Brothers, 1990). It has also been suggested that social cognition, including ToM, could have emerged to make possible cheating detection, and, perhaps more important for ancestral human societies, to reinforce cooperation (Brune, 2006). In chapter 3 we consider the question: how does a cooperative behavior based on ToM emerge out of interactions among constituent components in the brain and between components and the external environment?

A limited number of attempts have so far been made at the constructive approach to ToM characterized by the use of computational models for simulating its autonomous acquisition. Among them, there are only a few studies which investigate the underlying mechanism of the acquisition of the recursion level in a

ToM (Nagata et al., 2010; Noble et al., 2010; Takano and Arita, 2006; Yokoyama and Omori, 2009). However, functions of ToM in these studies are procedurally defined a priori by the designers. We focus on the emergence of a ToM without defining it a priori by modeling the brain at the functional level. To do this, we use a *Functional Parts Combination* (FPC) model (Ogawa and Omori, 2002; Omori and Ogawa, 2001), which regards the brain at a functional level as composed of a set of functional parts and activation signals specifying selectively activated patterns.

1.3 Description of the Thesis

The brief description of the thesis is as follows. The second chapter describes the evolution of the planning ability. The third chapter discusses the emergence of the ToM by modeling the brain at the functional level. The fourth chapter summarizes the thesis.

2 Evolution of the Planning Abilities

2.1 Introduction

Future-directed behavior can be seen in many animals as well as humans. For example, some hibernators store food for the coming winter just like humans who start building a shelter already in summer preparing for cold winter. So what is the difference in future-directed behavior between animals and humans? It has been said that animal behavior is instinctive but human behavior is flexible. *Mental time travel* is one of the capacities that provide increased behavioral flexibility of humans. Mental time travel is a term to refer to the faculty that allows humans to mentally project themselves backward in time to relive, or forward to prelive, events (Suddendorf & Corballis, 1997). The crucial selective advantage that mental time travel provides is the flexibility in novel situations and the versatility to develop and adopt strategic long-term plans to suit goals (Suddendorf & Corballis, 1997). In this chapter, we focus on the mental time travel into the future, especially the evolutionary aspect of the planning ability.

Our brains, which have high order functions such as planning, are energetically expensive. Although the human brain is only 2% of the body weight, it consumes about 20% of the total energy in the body (Clark & Sokkoloff, 1999). Furthermore, the prefrontal cortex, known to be critically involved in planning ability, accounts for nearly 30% of the cerebral cortical surface in humans (Brodmann, 1925). These facts raise the question of what are cognitive benefits to increased brain size. Natural selection has been considered as one of the most widely held mechanisms to explain the emergence of living creatures' complex characteristics. Evolutionary psychology has attempted to explain psychological traits as adaptations as the functional products of natural selection or sexual selection. Recent studies have indicated that ecological pressures drove the evolution of intelligence of human (Byrne, 1997; Darwin, 1871; Hill, 1982;

Osvath & Gärdenfors, 2005; Potts, 1998; Tooby & DeVore, 1987). They produce abilities to get a resource such as prey from the environment and to prevent the use from a predator as a resource (Geary, 2004). For example, with the global shift to cooler climate after 2.5 million years ago, much of southern and eastern Africa probably became more open and sparsely wooded, and it exposed the hominids to greater risk from predators and drove them into a cognitive niche (Tooby & DeVore, 1987).

Yet, it is difficult to explain why humans evolved such extraordinary cognitive competencies only by the ecological factors, considering that many other species hunt, occupy savanna habitats, endured the same climatic fluctuations, and so forth (Flinn et al, 2005). A different approach to the problem of the evolution of intelligence of human involves the consideration of the social aspect (Alexander, 1971, 1990; Brothers, 1990; Dunbar, 1998; Humphrey, 1976; Jolly, 1999). Alexander (1990) argued that it (evolution of the intellect) was rather the necessity of dealing continually with our fellow humans in social circumstances that became ever more complex and unpredictable as the human line evolved (pp. 4-7). Co-operating with other people is considered to be one of the most important factors to deal with our fellows in social circumstances. Furthermore, symbolic communication seems to be indispensable to co-operate smoothly with other individuals. Brinck & Gärdenfors (2003) traced the difference between the ways in which apes and humans co-operate due to differences in communicative abilities, claiming that there is a strong connection between the evolution of planning and symbolic communication. However, there is little known about the specific mechanisms that underlie it.

Considering all of the above factors, this chapter explores the dynamics inherent in the mechanism of the evolutionary acquisition of the planning abilities, focusing on the benefits of the planning and the costs of it. The first goal is to elucidate the environment which drove the evolution of planning ability. The second goal is to explore the dynamics inherent in the mechanism of evolution of the planning ability in the social circumstances. Our main method consists of a constructive approach which attempts to create not only a symbolic model of a living system, but also a symbolic living object (Moreno, 2002). Accordingly, our models are elaborated without direct and precise reference to empirical biological reality, and allow a new means of computational experimentation to enable us to discover the universal principles of living systems (Moreno, 2002). Next section explains a planner, task, architecture, and fitness of each agent. Section 2.3 shows the basic experiments, and section 2.4 describes the evolutionary experiments. Section 2.5 summarizes the chapter.

2.2 The Model

2.2.1 Planner - Beam Search

Gulz (1991) argued that an organism is planning its actions if it has a representation of a goal and a start situation and it is capable of generating a representation of partially ordered set of actions for itself for getting from start to goal. This criterion presupposes three distinct processes: (1) developing a plan, (2) remembering representations that have been developed, and (3) remembering the set of actions from start to goal. Some kinds of representational space in our mind such as a working memory make possible these processes. In the following model, we define the inherent planning parameter as an attribute value which corresponds to storage and processing capacity of the working memory system.

A beam search algorithm (Ney et al., 1992) is adopted as the planner of each agent. The beam search utilizes a heuristic value, *h*, to estimate the approximate steps from the focal state to the goal state, by which partial solutions are evaluated. It also uses a beam width, *B*, which specifies the number of states that are stored at each level of the breadth-first search. A *BEAM* is used to store the states that are to be expanded in the next loop of the algorithm. Also, a *hash table* is used to store states that have been visited.

At the process of the planning, initially, there is a start state in the *BEAM* and the *hash table*, respectively. Each time through the main loop of the algorithm, the planner expands states connected to the nodes in the *BEAM*, and adds the successor states to the *SET*, which stores all successors of the states in the *BEAM* at the current level, if they are not in the *hash table*, and then adds the best *B* states ordered by *h* from the *SET* to the *BEAM*. Note that if the high-priority states in the *SET* have the same heuristic value, some states are randomly chosen, and added to the *BEAM*. If the number of expansion reaches the inherent attribute value of the agent (termed *planning limit*), planner runs through the main loop and then *sub goal* is determined by selecting a state with the best *h* in the *hash table* other than the start state. If all states in the *hash table* have the same *h*, *sub goal* is randomly chosen from states in the *hash table*. Finally, solution is obtained tracing the path from a *sub goal* to the start state.

Table 2.1 shows an example trace of the algorithm on the state space in Fig. 2.1. As presented in Table 2.1, the more the number of times of expanding is, the deeper the search is. Also, agents who have long planning limit require large amounts of (1) storage capacity of the hash table and (2) processing power, such as attention control.

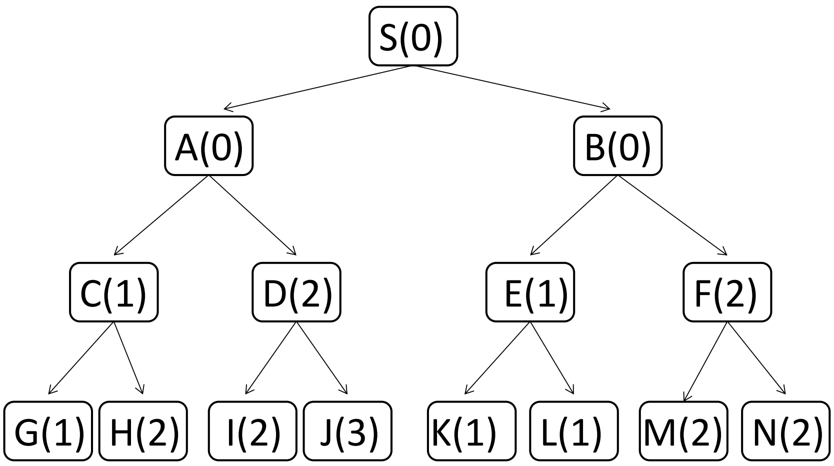


Figure 2.1: An example of the state space. Boxes with alphabets represent distinct states. S represents the start state. The numbers in boxes represent h.

Table 2.1: An example trace of the beam search algorithm on the state space in Fig. 2.1 when the B = 3. Each superscript represents the value of corresponding h. Each row shows the trace of the search when the planning limit of the agents is different. For example, agents with a planning limit of 2 choose D^2 as a sub goal. Also, agents with a planning limit of 3 choose D^2 or F^2 as a sub goal.

loop number	*planning limit*	*SET*	*BEAM*	*hash table*	*sub goal*
		{ }	$\{S^0\}$	$\{S^0\}$	{ }
1	1	$\{A^0,B^0\}$	$\{A^0,B^0\}$	$\{A^0,B^0,S^0\}$	$\{A^0$ or $B^0\}$
2	2	$\{C^1,D^2\}$	$\{B^0\}$	$\{C^1,D^2,A^0,B^0,S^0\}$	$\{D^2\}$
	3	$\{E^1,F^2,C^1,D^2\}$	$\{D^2,E^1,F^2\}$	$\{E^1,F^2,C^1,D^2,A^0,B^0,S^0\}$	$\{D^2$ or $F^2\}$
3	4	$\{I^2,J^3\}$	$\{E^1,F^2\}$	$\{I^2,J^3,E^1,F^2,C^1,D^2,A^0,B^0,S^0\}$	$\{J^3\}$
	5	$\{K^1,L^1,I^2,J^3\}$	$\{F^2\}$	$\{K^1,L^1,I^2,J^3,E^1,F^2,C^1,D^2,A^0,B^0,S^0\}$	$\{J^3\}$
	6	$\{M^2,N^2,K^1,L^1,I^2,J^3\}$	$\{J^3,I^2,N^2\}$	$\{M^2,N^2,K^1,L^1,I^2,J^3,E^1,F^2,C^1,D^2,A^0,B^0,S^0\}$	$\{J^3\}$

2.2.2 Task - Blocks World Problem

Planning is important especially when it is necessary to perform actions in the proper sequence to solve problems. We adopted the blocks world problem as a minimal task to deal with such a situation. A blocks world consists of a table with the size T, l rectangular blocks labeled $b_l(l = 1, \cdots, L)$, and a grip. The size of the table represents the maximum number of blocks that can be placed on the table. Each space of the table is labeled as $t_i(i = 1, \cdots, T)$. An agent is allowed to move a block to the top of another stack of blocks or to the empty space on the table by using a grip. A block can be moved only if there is no block on the top of it. In our model, if the table size is large, agents have many choices to move the block. Given the initial and target configurations of the blocks, the blocks world problem asks for a sequence of manipulation of the grip to achieve the target configuration with a smaller number of manipulations. In this study, we defined a target state as a configuration in which all blocks are stacked on a predetermined space in descending order as shown in Fig. 2.2. We define the heuristic value h representing the attainment level of the goal state as follows ($on(b_l, x)$: block b_l is on x; $\wedge$: logical symbol (and); $\neg$: logical symbol (not)).

$h = 5$: $on(b_1, t_3) \wedge on(b_2, b_1) \wedge on(b_3, b_2) \wedge on(b_4, b_3) \wedge on(b_5, b_4)$
$h = 4$: $on(b_1, t_3) \wedge on(b_2, b_1) \wedge on(b_3, b_2) \wedge on(b_4, b_3) \wedge \neg on(b_5, b_4)$
$h = 3$: $on(b_1, t_3) \wedge on(b_2, b_1) \wedge on(b_3, b_2) \wedge \neg on(b_4, b_3)$
$h = 2$: $on(b_1, t_3) \wedge on(b_2, b_1) \wedge \neg on(b_3, b_2)$
$h = 1$: $on(b_1, t_3) \wedge \neg on(b_2, b_1)$
$h = 0$: $\neg on(b_1, t_3)$

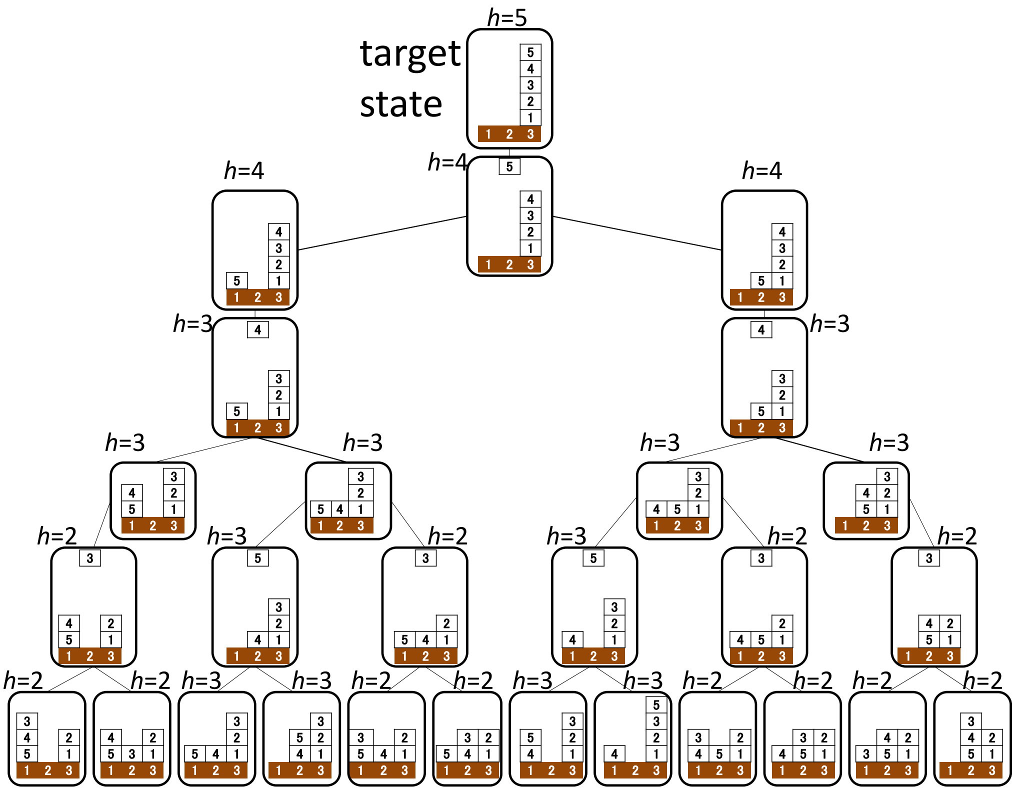
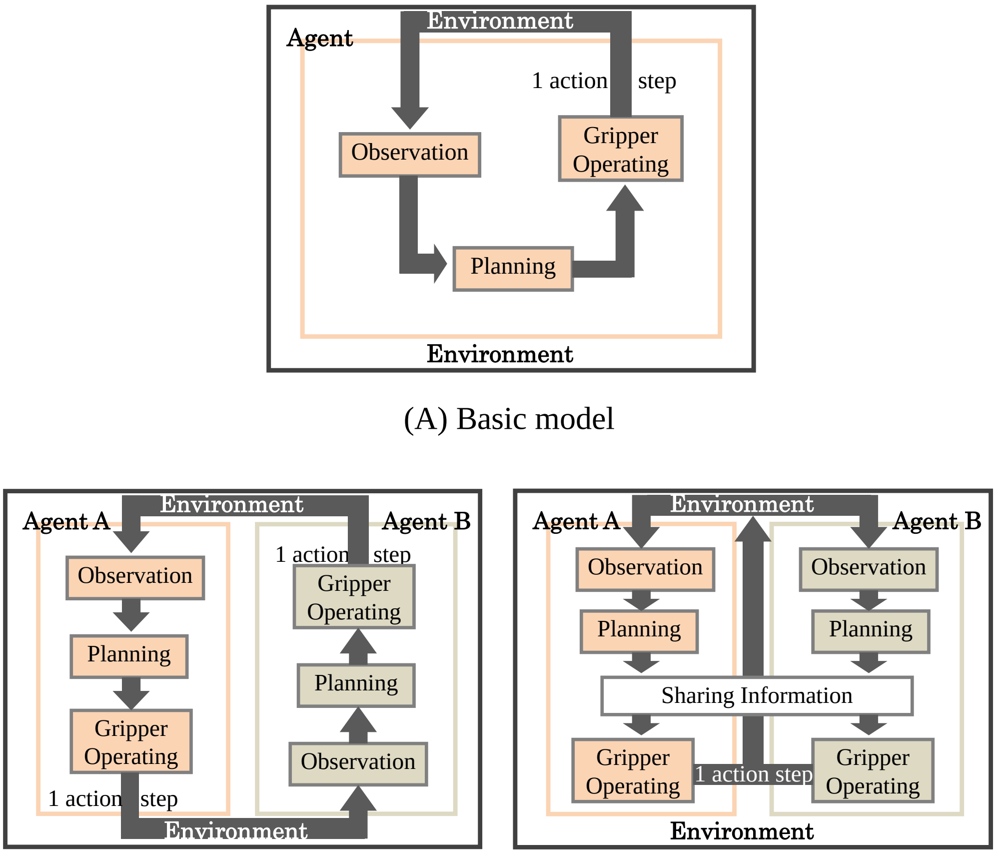


Figure 2.2: A part of the state space of blocks world problem (T=3). "h" represents the heuristic value corresponding to the each configuration.

2.2.3 Architecture

We constructed three models: basic model, nonshared model, and shared model. In the basic model (top in Fig. 2.3), an agent observes the present state at first, and passes it to the planner. Then, the planner makes a plan. Next, the agent moves a block once by the gripper. We define an action step, a, as a movement of a block by a gripper. Action steps are repeated until the target configuration is achieved or the number of performed action steps exceeds the upper limit a_{max}.



(A) Basic model

(B) Nonshared model

(C) Shared model

Figure 2.3: Architecture of an agent (top: basic model; bottom left: nonshared model; bottom right: shared model).

As to the nonshared and shared models, we assumed a situation in which two agents (agent A and agent B) participate in a collective task to reach the same goal. In the nonshared model (bottom left in Fig. 2.3), two agents interact by turn-taking: Agent A makes a plan, moves a block once by the gripper, and then agent B changes places with the partner A (makes a plan and moves a block once by the gripper). This cycle is repeated until the agents accomplish a goal, or exceeds an upper limit a_{max}.

In contrast, in the shared model (bottom right in Fig. 2.3), both agents make plan at the same time, and a plan which has the better heuristic value is selected. In the case in which the heuristic values of both agents are the same, the plan with a shorter sequence of actions is selected. When the length of both sequences are the same, either one is randomly selected. After that, both agents move a block once by the gripper. This cycle is repeated until the agents accomplish a goal, or exceeds an upper limit a_{max}. It is plausible to presume that information sharing during the plan selection process is based on symbolic communication.

2.2.4 Fitness

Each agent solves blocks world problem several times, and was evaluated by the fitness function *F*:

$$F = \frac{1}{T_{cost}}, \tag{2.1}$$

$$T_{cost} = w_a \times a_{cost} + w_p \times p_{cost} \qquad (w_a + w_p = 1), \tag{2.2}$$

$$a_{cost} = \frac{a - a_{min}}{a_{max} - a_{min}}, \tag{2.3}$$

$$p_{cost} = \frac{p - p_{min}}{p_{max} - p_{min}}, \tag{2.4}$$

where a is the average number of action steps that each agent performs to reach the target configuration among total trials of each agent, p is the planning limit of each agent, and a_{min}, a_{max}, p_{min}, and p_{max} are fixed numbers. Also, w_a

is the weight to the action cost (a_{cost}), and w_p is the weight to the planning cost (p_{cost}). Equation (2.1) suggests that the greater the action steps or the planning limit is, the lower the fitness is. The balance between the cost of action and planning is determined by w_p and w_a.

2.3 Basic Experiments

2.3.1 Experimental Setup

The experiments in this chapter focused on the two parameters controlling the difficulty of the problem, thereby investigating the conditions of the environment for the planning ability to evolve: The depth of the optimal solution (*D*) and the size of the table (*T*). The depth was defined as the shortest path from start to goal. Also, as the size of the table (as we mentioned before) becomes large, the number of optimal paths on the state space increases because agents have many choices when they move a block in the case the table size is large.

We conducted basic experiments (in which the planning limit of agents was not evolved but fixed) to find how obtained solution is influenced by the difference in planning limits, problem difficulty or collective manner. The experiments were conducted with 3 different *D* values varied by changing start configurations (goal configuration was fixed) while fixing *T*, and the ones with 4 different *T* values while fixing *D* (both start and goal configurations were not changed) as shown in Table 2.2.

Table 2.2: Experimental setup of the basic experiment.

L (number of blocks)	5						
a_{max} (upper limit)	500						
B (beam width)	7						
Number of trial run	100						
T (table size)	4			3	4	5	6
D (depth of the problem)	12	14	16	18			

2.3.2 Results

Fig. 2.4 shows the average action steps to solve the problem in various settings of the *T*, *D* and planning limit in the basic model. It is shown that the solution became worse (actions steps became larger) when *T* was smaller and *D* was higher. This is because the problem became more difficult to solve in those situations. This tendency was stronger when planning limit was smaller. Especially when the planning limit was the minimum, the agent took a random action because the planner explored the states with the same heuristic value next to the present state.

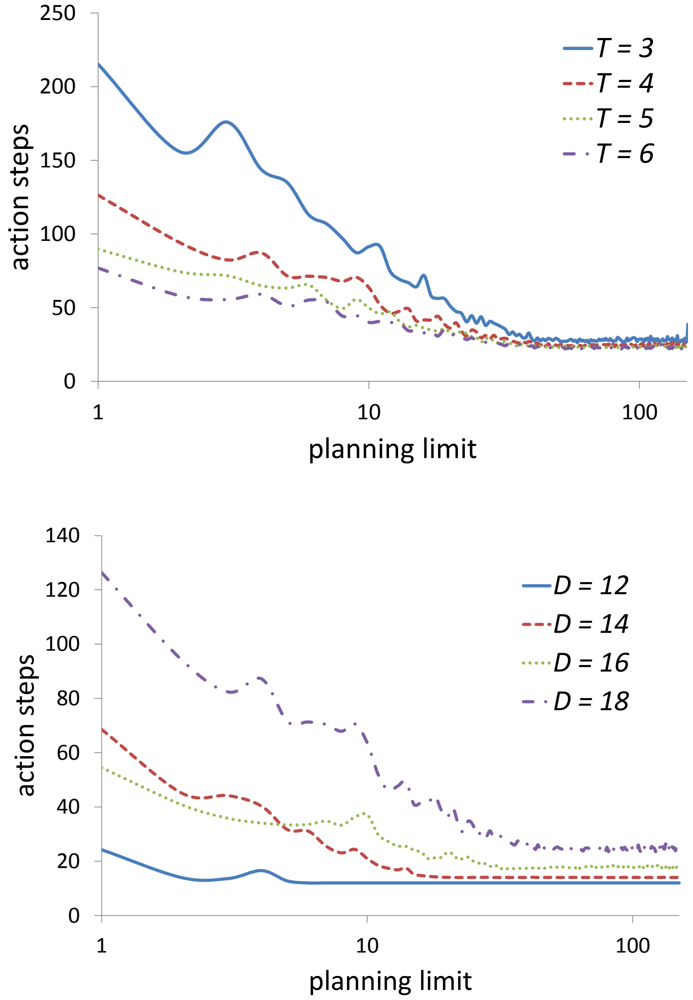


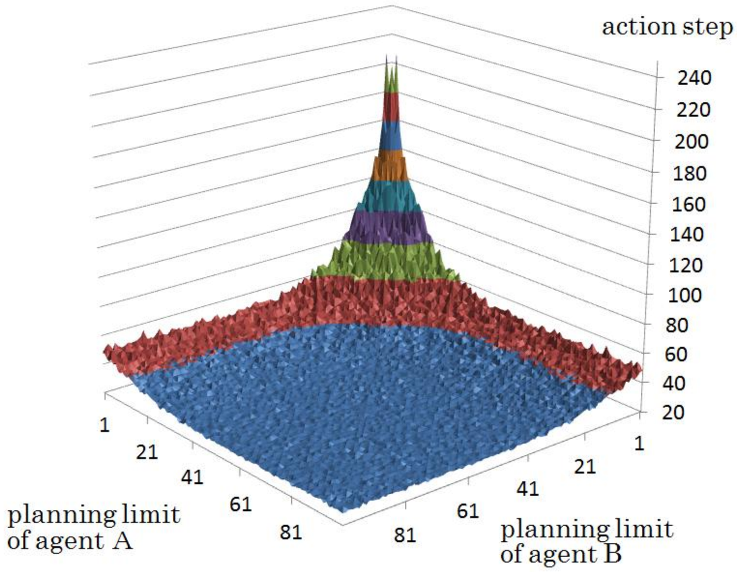
Figure 2.4: Average action steps to solve a problem in various settings of the table size (T), the depth of the problem (D), and the planning limit in the basic model.

On the other hand, when the planning limit was high, the agent took a proper action toward the goal because the planner explored near-goal states far from the present state. Simulation showed the same tendency both in the nonshared and shared models.

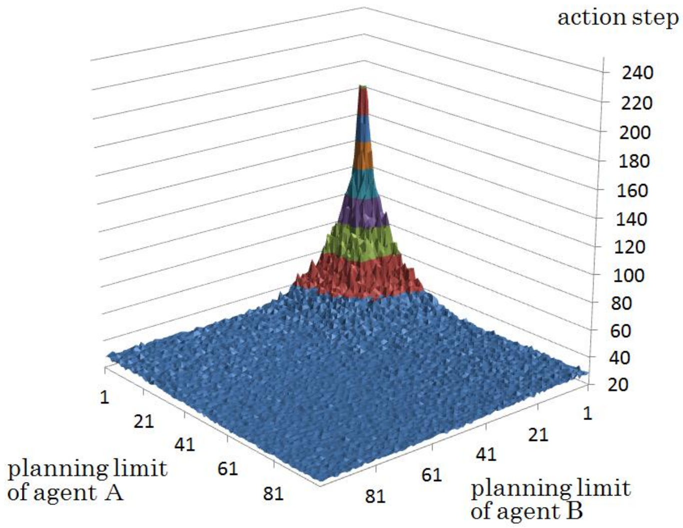
Fig. 2.5 shows the action steps averaged over 100 trials of the task evaluation when varying the planning limits of two agents (top: nonshared model; bottom: shared model). The x and y axes correspond to the planning limit of agent A and that of agent B, and the z axis represents the average action step. We can find that the effects of the planning limits of both agents on the obtained solution were complementary. In other words, it was possible to decrease action steps when the planning limit of either one was long, even if that of the other was very short.

Fig. 2.6 shows the average action steps when two agents took the same planning limit. As shown in Fig. 2.6, the quality of the solution of the shared model was more improved than that of the nonshared and basic model for the following reason. At the process of the planning, intermediate states in the *BEAM* and *sub goal*s are randomly chosen under some conditions, and it varied in plans. Since a plan with a shorter sequence of actions is selected if the heuristic value of both agents is the same in the shared model, agents in this model could behave more efficiently by comparing both plans.

Fig. 2.7 shows the fitness landscape using the data of the action steps in Fig. 2.5 where w_p was 0.1 (top: nonshared model; bottom: shared model). The x and y axes correspond to the planning limit of agent A and that of agent B, and the z axis represents the fitness of agent B. In the nonshared model (Fig. 2.7 - top), agents who had the middle planning limit (in the range of about 20 to 40) could get the highest fitness when that of the other was long (in the range of about 40 to 100). On the other hand, in the shared model (Fig. 2.7 - bottom), agents who had extremely a short planning limit (in the range of about 1 to 10) could get the highest fitness when that of the other was long (in the range of about 40 to 100) for the following reason. In the shared model, agents who have extremely a short planning limit can decrease action step when that of the other was long because solutions by the agent who has the longer planning limit tend to be adopted in almost all trials. Therefore, they can get the highest fitness because of the low cost for planning.



(A) Nonshared model



(B) Shared model

Figure 2.5: Action steps averaged over 100 trials of the task evaluation when varying the planning limits of two agents where F=18 and T=3 (top: nonshared model; bottom: shared model).

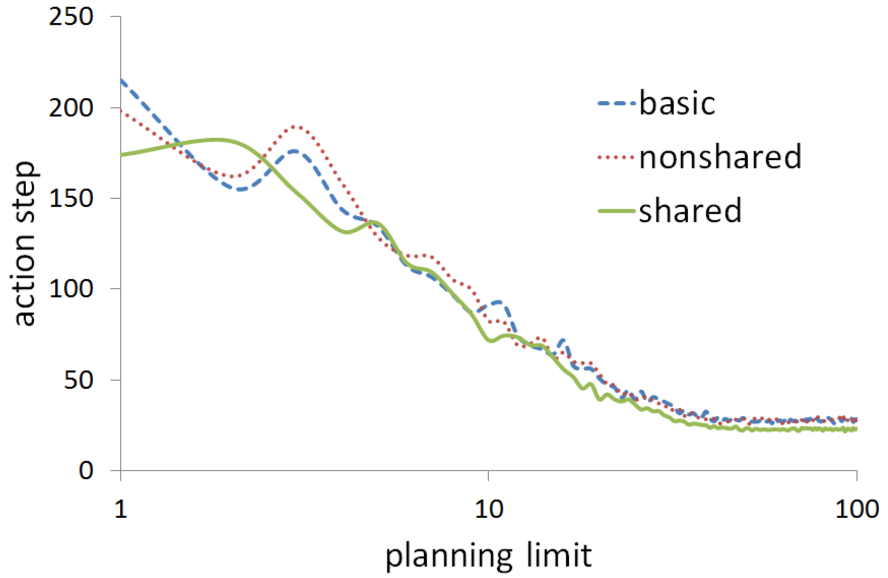
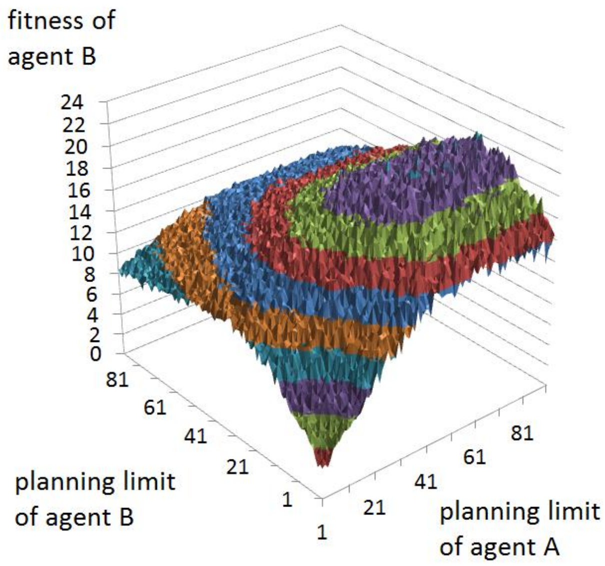
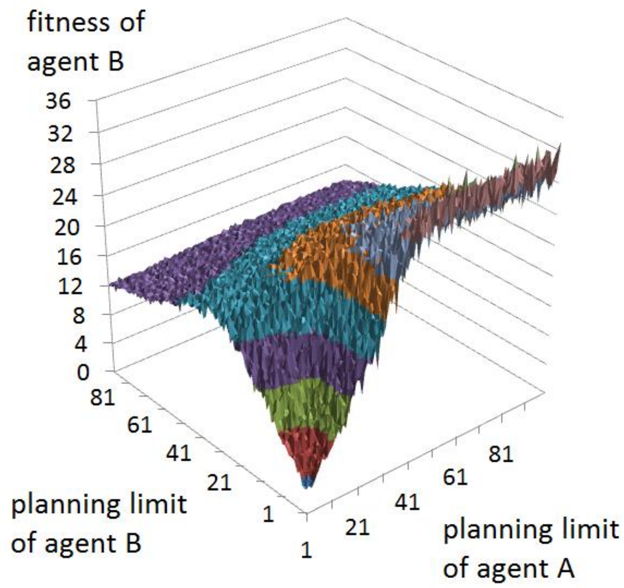


Figure 2.6: Average action steps of 100 trials of the task evaluation when two agents took the same planning limit where F=18 and T=3.



(A) Nonshared model



(B) Shared model

Figure 2.7: Fitness landscape of the agent B where action steps were averaged over 100 trials of the task evaluation when varying the planning limits of two agents where F=18, T=3, and w_p*=0.1 (top: nonshared model; bottom: shared model).*

2.4 Evolutionary Experiments

2.4.1 Experimental Setup

We conducted simulations in which the planning limit of agents was evolved by using a genetic algorithm. A chromosome was represented by integer encoding, which determines the planning limit of each agent. We first created N individuals whose planning limits were randomly selected from 1 to 3. As to the basic model each agent solved blocks world problem H times. As to the nonshared and shared model, every pair of agents solved the problem in a round robin manner. Then, action steps were averaged over those games, and agents were evaluated by the fitness function (2.1).

The offspring in the next generation were selected by the ranking selection as follows. The selection probability p_i is defined by using the scaled fitness $f_i^{'}$ as:

$$p_i = \frac{f_i^{'}}{\sum_{j=1}^{N} f_j^{'}}. \tag{2.5}$$

Here, $f_i^{'}$ is defined as

$$f_i^{'} = (N - R_i + 1)^2, \tag{2.6}$$

where R_i is the fitness rank of individual i. Then, each gene of all offspring was mutated with a probability P. In the phase of mutation, a random integer digit m was generated from a uniform distribution between -M to +M, and added to the original genetic value.

Table 2.3: Experimental setup of the evolutionary experiment.

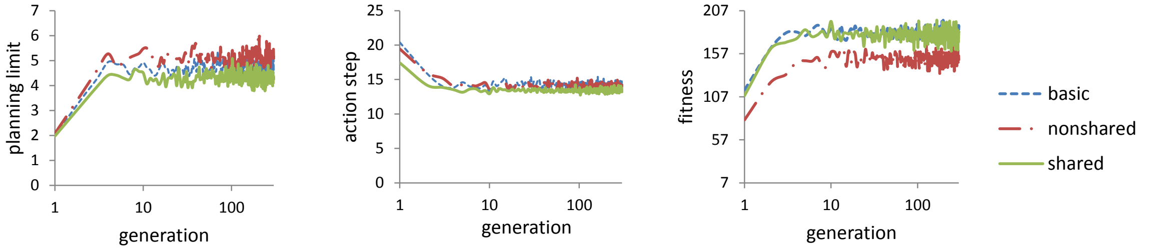
L (number of blocks)	5						
B (beam width)	7						
a_{max}	500						
a_{min}	11						
p_{max}	200						
p_{min}	0						
H (repeat number of times)	5						
N (population size)	20						
P (mutation rate)	0.5						
M (mutation range)	5						
generation	300						
Number of trial run	10						
T (table size)	4			3	4	5	6
D (depth of the problem)	12	14	16	18			

2.4.2 Results

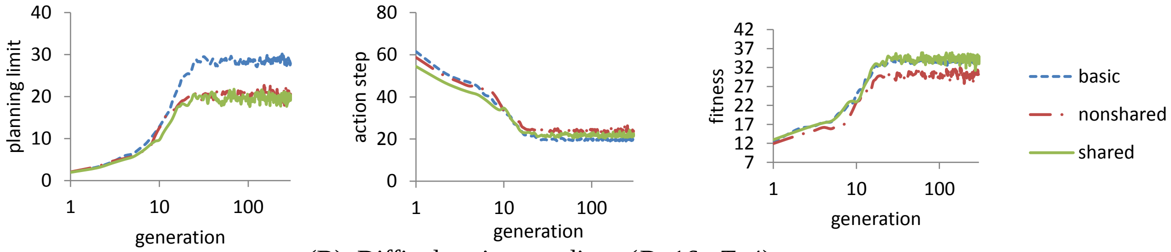
We conducted evolutionary experiments using parameters of Table 2.3 where w_p was 0.1, 0.3, and 0.4. Fig. 2.8 shows the difference in transition of the fitness, action steps, and the average planning limit among individuals between basic model, nonshared model, and shared model when w_p=0.1. Top row (A), middle row (B) and bottom row (C) show the results of the easiest task (D=12, T=4), an intermediate task (D=16, T=4), and the most difficult task (D=18, T=3), respectively. We can find that the action steps decreased and the planning limit increased through the course of evolution. Also, long planning limits were emerged when the difficulty of the problem was intermediate (Fig. 2.8 - (B)) or high (Fig. 2.8 - (C)) because agents could minimize action steps as a reward for increasing the planning limit. On the other hand, long planning limits were not so emerged when the difficulty of the problem was low (Fig. 2.8 - (A)) because action steps only slightly changed even if the planning limit increased. Simulation also showed the same tendency when w_p=0.3 and 0.4 as shown in Fig. 2.9 which represents the results of the average planning limits between 200 to 300 generations in the basic model, nonshared model, and shared model when the difficulty of the given problem was changed.

Fig. 2.10 shows the transition of the distribution of the planning limits in the population on a certain trial of the basic, nonshared, and shared models when the difficulty of the problem was high. As shown in Fig. 2.10, agents who had even the short planning limit (in the range of 10 to 29) could exist in the collective (nonshared and shared) models. Also, the average planning limit in the collective models was less than that in the basic model when the difficulty of the problem was intermediate (left in Fig. 2.8 - (B)) or high (left in Fig. 2.8 - (C)). This is because the effects of the planning limit of both agents on the obtained solution were complementary, and it was possible to decrease action steps when the planning limit of either one was long, even if that of the other was short (Fig. 2.7). Therefore, in a case in which agents who had a long planning limit occupied a major part of population, agents who had a short planning limit could enter the population because they could get relatively great fitness because of the low cost for planning. This situation was equivalent to the tragedy of the commons (Hardin, 1968). It explains the reason why the planning ability was difficult to emerge in the collective situation. The reason why difference in planning limits of the most difficult task between the basic and collective (nonshared and shared) model was smaller than that of the intermediate task is that the merit of the planning limit would slightly weaken the force of free rider problem.

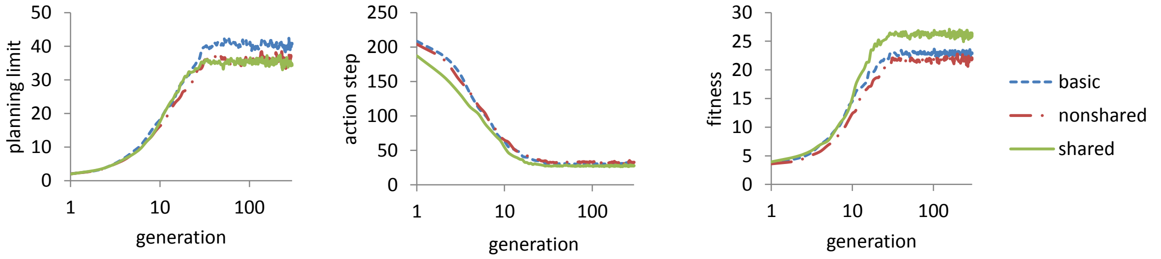
The notable point is that planning ability equally evolved both in the shared and nonshared model even though the free rider problem tends to be more serious in the shared model as shown in Fig. 2.7. This may be because the optimization of the planning limit proceeded for the following reasons. First, sharing information improved a quality of the solution. Second, planning of the shared model have a greater tendency to have a positive effect on the solution than that of the nonshared model when the own planning limit is longer than other agent. As a result, planning of the shared model equally evolved to that of the nonshared model, and fitness of the shared model was higher than that of nonshared model.



(A) Difficulty: low (D=12, T=4)



(B) Difficulty: intermediate (D=16, T=4)



(C) Difficulty: high (D=18, T=3)

Figure 2.8: Difference in transition of the fitness, action steps, and planning limits between the basic model, the nonshared model, and the shared model when the difficulty of the given problem was changed (w_p=0.1).

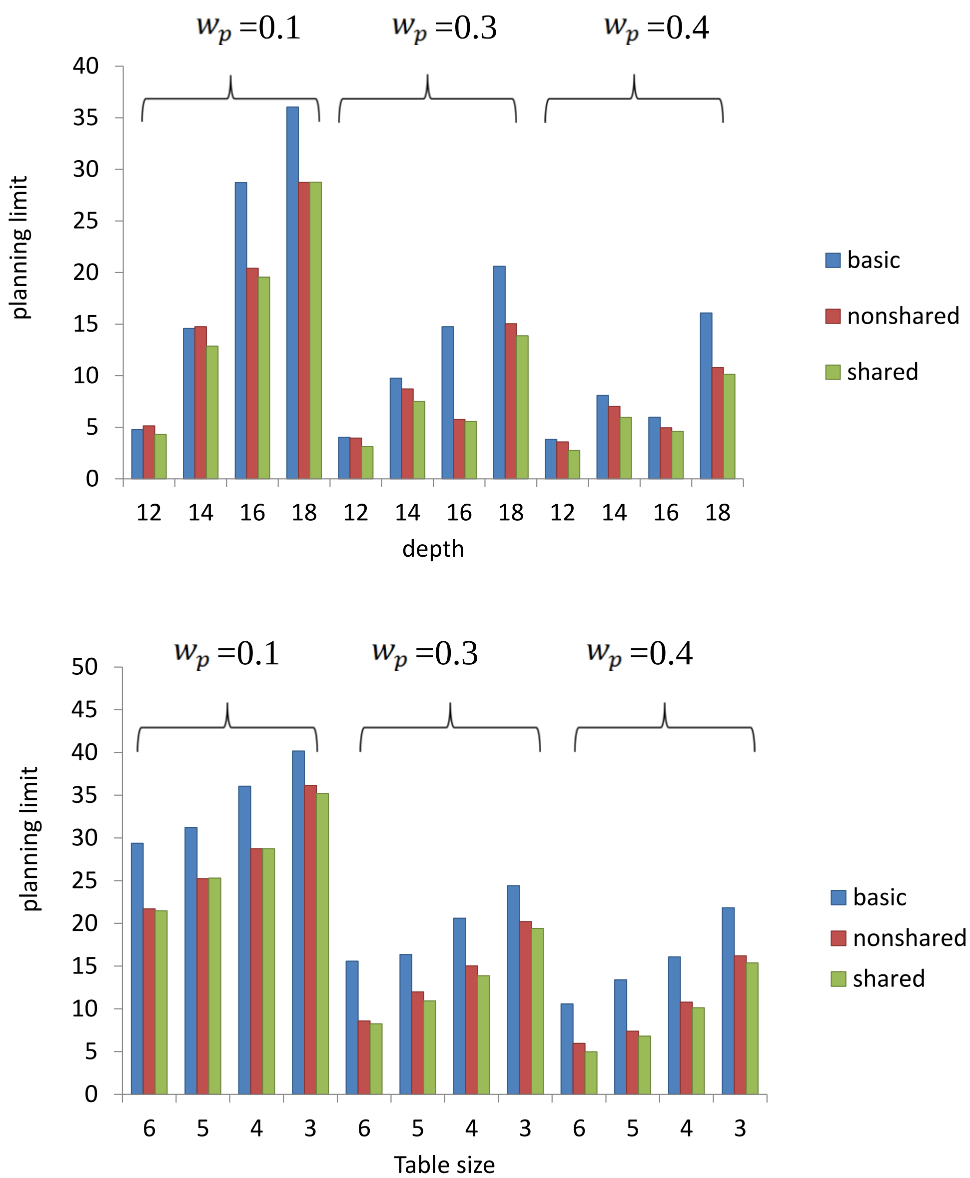
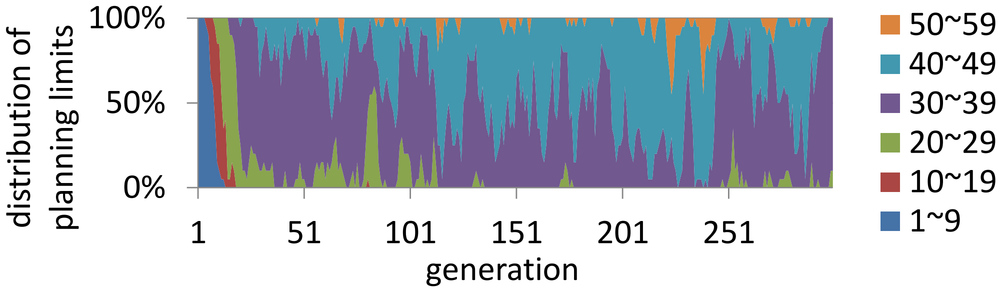
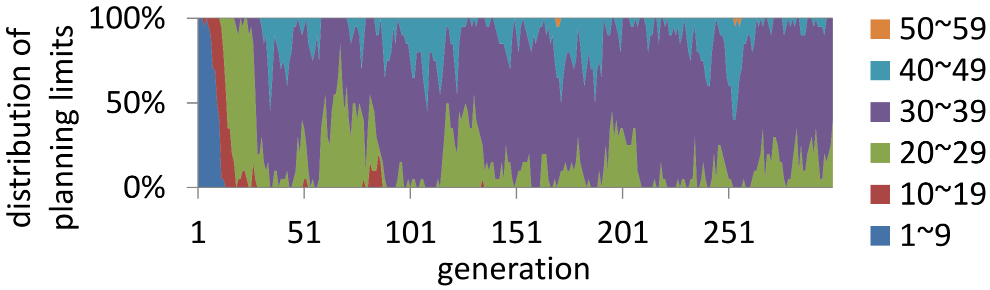


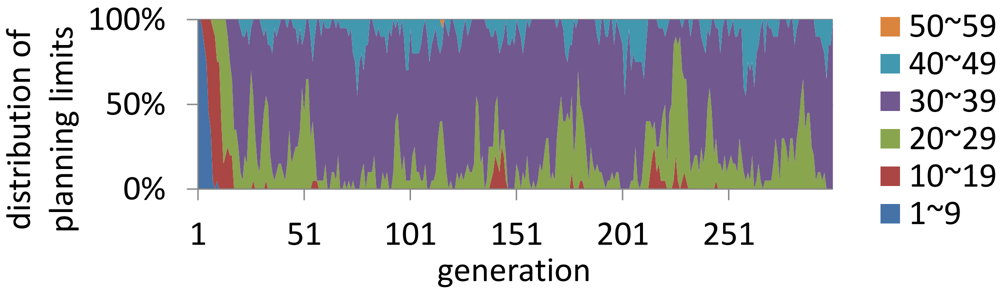
Figure 2.9: Average planning limits between 200 to 300 generation in the basic model, the nonshared model, and the shared model when the difficulty of the given problem was changed.



(A) Basic model (D=18, T=3, w_p=0.1)



(B) Nonshared model (D=18, T=3, w_p=0.1)



(C) Shared model (D=18, T=3, w_p=0.1)

Figure 2.10: Transition of the distribution of the planning limit in the population on a certain trial.

2.5 Conclusion and Discussion

In this chapter, we investigated the mechanism of the evolution of the planning abilities by focusing on the benefits and the costs of the planning. Table 2.4 summarizes the results of the evolutionary experiments in the basic model, nonshared model and shared model, where the difficulty of the problem was changed. The size of the circles corresponds to the relative length of the evolved planning limits, action steps and merit of the sharing information in several settings. As shown in Table 2.4, simulation clarified that the longer planning limit emerged as the given problem became difficult to solve both in independent and collective situations. So what does the difficulty of the problem mean in the evolution of the human intelligence? It has been said that ice sheets started to grow in the northern parts of the world, and Africa experienced deforestation and expanding savannas. These conditions in savannah might force the hominid to use a wide variety of food sources which were more transient and scattered than the predominantly vegetarian food sources (Bickerton, 2002). It might work as a selective pressure for more efficient feeding, and thus an increasing need for sophisticated tool use by planning might be selected (Byrne, 1997). Yet, if the individuals face more difficult problems such as making complicated stone tools, making a tent, farming, or stock raising, working together seems to be important to accomplish a goal efficiently.

Table 2.4: Conceptual diagram of the evolution of the planning abilities given by the results of the evolutionary experiments.

	difficulty: low			difficulty: intermediate			difficulty: high		
	basic model	nonshared model	shared model	basic model	nonshared model	shared model	basic model	nonshared model	shared model
evolved planning limit									
action step									
merit of the sharing information									

The simulation results indicated that when the problem was difficult, planning ability was difficult to evolve in the collective situation because there was a conflict between personal and collective interest (Table 2.4). So, how could hominids climb the steep cost gradient? The results also clarified that planning ability equally evolved both in the situations where individuals shared information and did not share information even though the free rider problem tended to be more serious in the former situation. Also, fitness of the situation where individuals share information was higher than that of individuals did not share information because of improving a quality of the solution. Considering all of the above factors, we can present the following scenario as: (1) First, a selective pressure for more efficient feeding in savannah made the use of prospective cognition that is the skill to plan for future events and needs, beneficial; (2) As the problem became more difficult, increasing need for collectively work would be selected, however; a cost of thinking might be serious at the same time; (3) The select of symbolic communication might be favored because it was an efficient way of solving problems. This result implied that there is a connection between evolution of the planning and symbolic communication.

3 Emergence of a Theory of Mind

3.1 Introduction

Cooperative behaviors are one of the fundamental processes which form societies in animals. We will use the term *cooperation* as: cooperation is a type of interaction involving two or more agents who are (1) trying to achieve one or more common purposes, and (2) making proper coordination between their behaviors (Namatame, 2004). It is found everywhere in the nature. For example, there are division of labors in insect societies such as ants and bees. However, the action decision processes of each individual are very simple and instinctive. On the other hand, there are complex and dynamic action decision processes in cooperation of humans: interactively determining own behavior in accordance with the estimated intention of others. In general understanding of others as having intentional states such as beliefs and desires is called *Theory of Mind* (ToM) (Premack and Woodruff, 1978). The aim of this chapter is twofold. The first is to investigate how cooperative behaviors based on ToM emerge through autonomous developmental and evolutionary processes by modeling the brain at the functional level. The second is to apply our model as a framework for the mind system of human-like robots with the capacity of estimating.

A specific type of cells called mirror neurons have been found in the premotor cortex of monkeys, an area that is possibly homologous to the Broca area in humans (Gallese and Goldman, 1998). These neurons fire both while monkeys or human act and while they observe the same action performed by another. It has been suggested that mirror neurons might be part of, or a precursor to, a more general mind-reading ability. In addition to this, the discovery of mirror neurons has an impact on the scientist's views on the brain. When scientists referred to the premotor cortex in the past, they assumed that each function of the brain was handled under a paradigm of functional localization. However, to have reaction characteristics such like a mirror neuron system, it has to integrate visual and motor information at least, complex functions that remained unexplained by earlier paradigms.

By the subsequent investigations, it is now widely accepted to regard the brain as a whole system integrating the multiple component parts. The *Functional Parts Combination* (FPC) model (Ogawa and Omori, 2002; Omori and Ogawa, 2001), which will be explained later, is a general model of the brain at the functional level which describes such a system. We began a study on how ToM has been shaped through the autonomous developmental process, assuming a functional model of the brain as an FPC model. A limited number of attempts have so far been made at the constructive approach to ToM characterized by the use of computational models for simulating its autonomous acquisition. Among them, there are only a few studies which investigate the underlying mechanism of acquisition of the recursion level in a ToM (Nagata et al., 2010; Noble et al., 2010; Takano and Arita, 2006; Yokoyama and Omori, 2009). However, functions of ToM in these studies are procedurally defined a priori by the designers.

We focus on the emergence of a ToM without defining it a priori by adopting an FPC model. In this chapter, we propose a computational framework for investigating the emergence of a ToM based on adaptation in both evolutionary and individual-learning time scales. As a first step, this chapter also reports on the results of the computer simulation which demonstrates an acquisition of the activated patterns of the functional parts for processing ToM through learning. Next section shows three hypotheses postulated in our research for investigating the emergence of ToM.

Section 3.3 explains a functional model of the brain and Section 3.4 illustrates a task and an intention estimation model. Section 3.5 and 3.6 shows the experiments and Section 3.7 summarizes the chapter.

3.2 Three Hypotheses for Investigating the Emergence of Theory of Mind

For our purposes mentioned above, we introduce following three hypotheses.

The first hypothesis is that brain is a complex system composed of multiple functional modules and dynamically searches a combination of the modules depending on the task (FPC model). Recent studies in the neuroscience have been elucidated that each cerebral cortical area has a different role and is selectively activated depending on the task (Liu et al., 1996; Toga and Maziotta, 2000). These facts imply that each brain areas is assigned a different function, and there is a system which combines a set of appropriate functional parts depending of the task. Meanwhile, when someone is confronted with the unprecedented tasks, multiple areas in the brain are activated simultaneously, and after a while a set of modules corresponding to the task is settled gradually. Therefore, a combination of the modules assumes processes based on the some kind of search. Omori and Ogawa (2002) conducted agent-based simulations based on the FPC model and showed that the agent using it acquires learning strategies suitable for the multiple tasks of navigation problem. If the task is dynamic such as social interactions supposing in this chapter, behaving swiftly depending on others would be critical because behavioral strategies of the agent often change depending on one's own behavior.

The second hypothesis is that humans modulate our cognitive state dynamically depending on recognition of interactive agents. This view has much in common with the *conflict monitoring hypothesis* (Botvinick et al., 2004) where the anterior cingulate cortex (ACC) is believed to be the area which detects the occurrence of conflicts in information processing, and thus triggers strategic adjustments in cognitive control. Takahashi et al. (2008) proposed a computational model, based on the conflict monitoring hypothesis, which consists of change detection and state space switching evoked by the change of environmental nature. In order to realize a smooth cooperative behavior, we adopted the attention system of the Takahashi et al. (2008) in order to dynamically modulate the cognitive state in the FPC model.

The third hypothesis is a *simulation-theory* (Gallese and Goldman, 1998; Gordon, 1995) which assumes that human can understand other minds of others by means of a simulation process such that we use our own mental mechanisms to read the mind of others. The alternative hypothesis, often-quoted, is a *theory-theory* (Carruthers, 1996) which assumes that humans attribute mental states to others using theoretical considerations involving a set of concepts (beliefs, desires, and so forth) and principles about how these concepts interact. Although both hypotheses can be consistent with the view that the neural basis of understanding other minds is localized in the frontal cortex (Toga, 2000), we constructed a computational model under the simulation-theory. The main reason is that the discovery of mirror neurons accords well with simulation-theory. In addition, Gordon (1995) suggests inferiority of the theory-theory from the perspective of 'economy storage': It would be extravagant if we had to store in memory theoretical considerations on the numerous relationships between stimuli and mental states, mental states and other mental states, and mental states and behavior. On the other hand, simulation-theory does not need such effort because other people's mental states are represented merely by adopting own perspective. From a developer's perspective, this problem is considered to be critical; i.e., effective and efficient model creation reduces development costs of virtual training systems (Harbers, 2009).

3.3 Functional Model of the Brain

We propose a functional model of the brain based on above hypotheses (Fig. 3.1). There are modules M_i in the brain, which constitute a module network. A set of modules in the network are activated by a set of activation signals A which is represented as a vector of binary values 0 and 1: $A = (a_{0,} \cdots, a_i, \cdots, a_{k-1})$, where k is the number of modules, and a_i is an activation signal for module M_i. Also, some links in the network are deactivated by a set of link-deactivation signals D which is represented as a vector of binary values 0 and 1: $D = (d_{0,} \cdots, d_i, \cdots, d_{m-1})$, where m is the number of links, and d_i is a link-deactivation signal for link l_i. In the sub-network, parallel computation is controlled based on the simple parallel control flow paradigm (Trealeven et al., 1982) as follows.

- The execution starts from the sensory input.
- All data are transferred indirectly between modules via updatable memory cells which are initialized before each sensory input.
- In a case that all input links receive a control token the activated modules begin its computation while the non activated modules do not execute it.

- Then activated modules output tokens at all output links. However, the modules whose input links are not connected do not output tokens, regardless of whether or not the modules are activated.

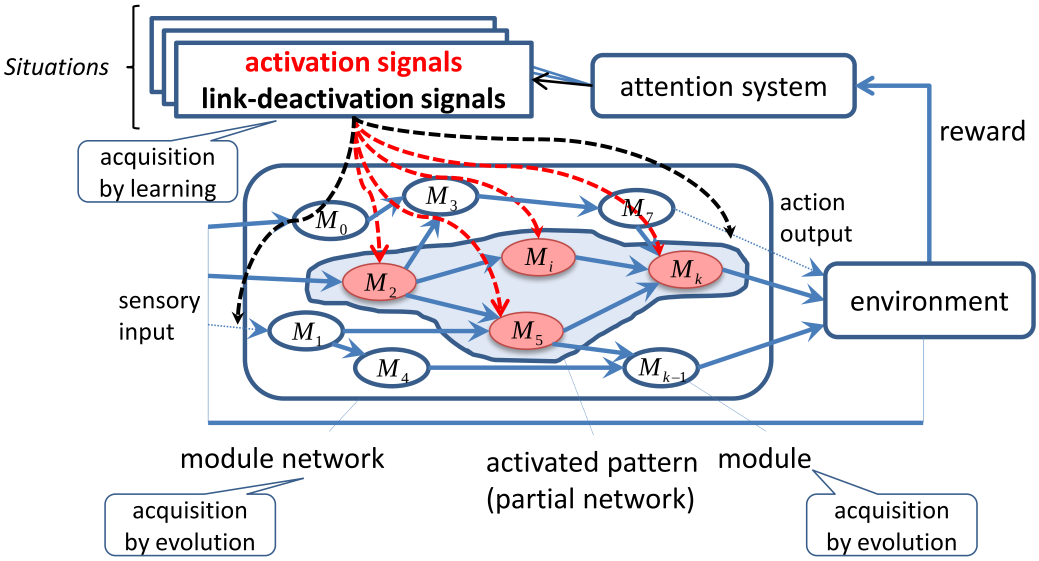


Figure 3.1: Functional model of the brain.

The brain searches for a set of activation and link-deactivation signals: *Solution* = (*A*, *D*) responding to the current task (or social situation). The searching system alternates between the two modes: searching for a new *Solution* and switching among the acquired *Solution*s, evoked by the change of environmental nature (Takahashi et al., 2008) (Fig. 3.2). In both modes, the evaluation of the pattern activated at each period is updated per unit time:

$$E_{p+1} = \alpha_e \cdot reward + (1 - \alpha_e) \cdot E_p, \tag{3.1}$$

where α_e is a parameter which coordinates the update rate and *reward* is an evaluation of the present partial network. The threshold is also updated per unit time:

$$T_{p+1} = \alpha_t \cdot T\_max + (1 - \alpha_t) \cdot T_p, \tag{3.2}$$

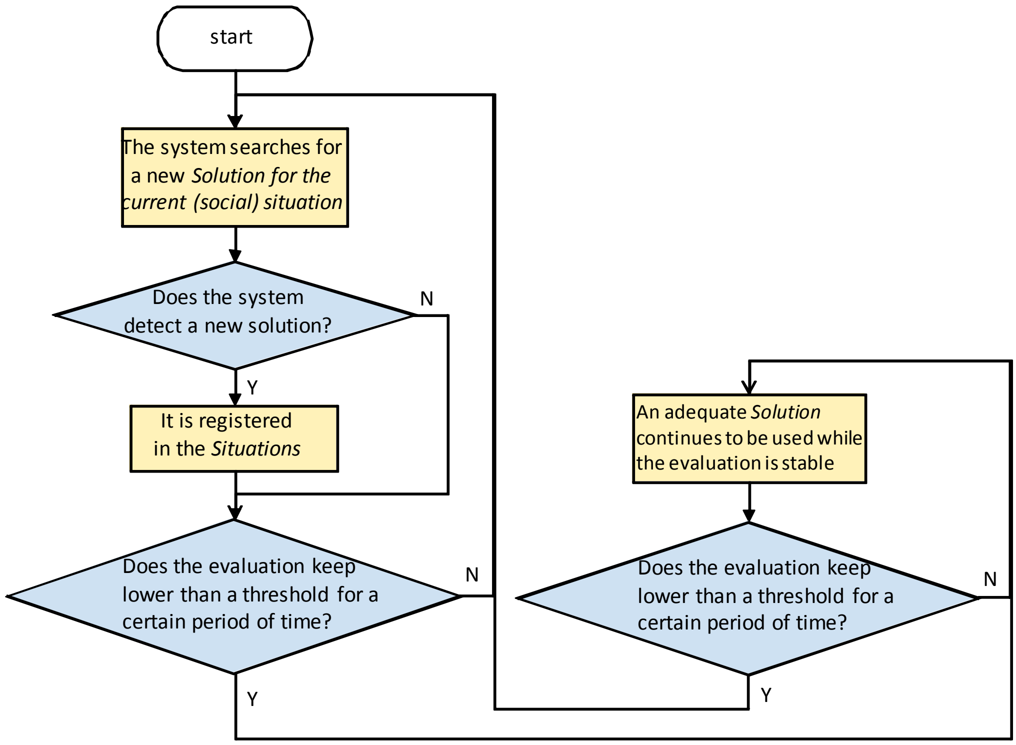


Figure 3.2: Algorithm of the searching system.

where α_t is a parameter which coordinates the update rate and *T_max* corresponds to a threshold value of the model. In searching mode, it searches for a new *Solution* for the current (social) situation. When a solution which forms a new partial network containing sensory input and action output is obtained, it is registered in the *Situations*. The system changes into switching mode when the evaluation keeps lower than a threshold for a certain period of time. In switching mode, an adequate *Solution* continues to be used while the evaluation is stable. However, if there are no suited solutions in the *Situations* for the current environment, which means that it is a novel environment, the system returns into searching mode. The sensitivity to change of the environmental nature is regulated by *T_max* in the expression (3.2). In order to prevent the frequent mode change, T_p is set to *T_min* after the mode change.

From the viewpoint of emergence, it is essential to discuss how to construct this functional structure in a self-organizing fashion. The module network and constituent modules can be acquired by evolution (e.g. Genetic Network

Programming (Hirasawa et al., 2001) and Genetic Programming (Koza, 1992), respectively). The set of activation and link-deactivation signals can be acquired by learning (e.g. Tabu Search (Glover and Laguna, 1997) or Simulated Annealing (Kirkpatrick et al., 1983)).

3.4 Task and Intention Estimation Model

As a first step, this chapter focused on the emergence of activation signals for forming ToM sub-networks with a specific recursion level to achieve cooperative behavior in a hunter task. The discussion on the emergence of modules network and the constituent modules is outside the scope of this chapter and we assumed that they had been acquired.

3.4.1 Hunter Task

There are two hunters and two prey in a 20×20 a two-dimensional grid folded to a torus (Fig. 3.3). Each hunter moves one cell per *step* to the left, right, up or down, or stays in the current cell according to its own strategy, while each prey moves one cell per *step* stochastically (right; 40 %, up; 20 %, or stop; 40 %). When starting the task, all 4 agents are randomly located in the grid, and each hunter selects the closer prey as an initial target. Each *episode* ends when each hunter captures the different prey or the number of time steps exceeds the upper limit $step_{max}$.

The task is solved several times using the current *Solution* in each *period*. Averaged time steps to solve the task, which is used to evaluate the solution:

$$reward = 100 / step. \tag{3.3}$$

The *reward* is shared between cooperators, and then E_p (expression (3.1)) and T_p (expression (3.2)) are updated.

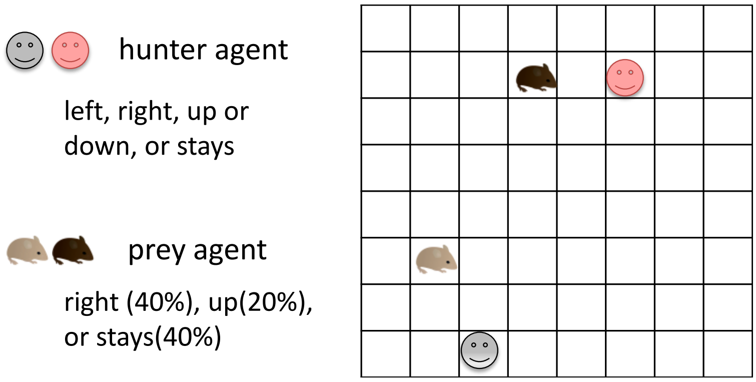


Figure 3.3: Hunter task.

3.4.2 Intention Estimation Model

The agents require an intention estimation of the other for solving the hunter task efficiently, in the sense that agents have to estimate the prey that the other agent tries to capture and chase the prey different from others'. As an intention estimation model, we assumed that humans estimate the intention of others by simulating it based on their own action-selection process as if they were in the same situation (simulation-theory). More formally, the intention of each agent is assumed to be determined by the three variables: goal G, state s and action a (Yokoyama and Omori, 2009). If the agents identify two of its three variables, the remaining one can be estimated through the acquired own experience and knowledge expressed as a joint probability distribution: $P(a, s, G)$. For example, the intention of others G can be estimated by using the conditional probability distribution: $P(G|\, s, a)$. In a similar way, the state s and action a of the others can be estimated by using $P(s|\, a, G)$ and $P(a|\, s, G)$, respectively. However, in a condition where agents mutually estimate the intentions of others, the cooperation performance is unsatisfactory (Nagata et al., 2010). Therefore, agents might have to flexibly change own action decision strategies depending on others. Considering the above factors, we constructed the following strategies based on a Dennett's *intentional stance* (Dennett, 1987) (Fig. 3.4):

(1) Agent at level 0 takes action a_s based on own state s_s and goal G_s without considering behavior of others:

$$a_s = argmax_a P(a|s_s, G_s). \tag{3.4}$$

(2) Agent at level 1 estimates the intention of others G_o based on other's state s_o and action a_o, by assuming that others would be at level 0:

$$G_o = argmax_G P(G|s_o, a_o), \tag{3.5}$$

and takes action based on it:

$$a_s = argmax_a P(a|s_s, G_s), \tag{3.6}$$

where G_s satisfies a cooperated condition. For this task $G_s \neq G_o$ (i.e. own goal differs from estimated other's goal) is simply a condition of cooperation.

(3) Agent at level 2 estimates the own intention which is estimated by others $\tilde{G}_s$ by assuming that others would be at level 1:

$$\tilde{G}_s = argmax_G P(G|s_s, a_s), \tag{3.7}$$

and takes action based on it:

$$a_s = argmax_a P\left(a \middle| s_s, \tilde{G}_s\right). \tag{3.8}$$

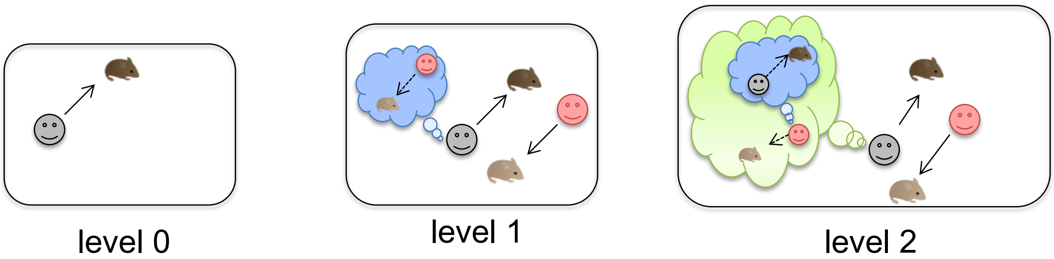


Figure 3.4: Higher-order estimation of other's intention in terms of intentional stance developed by Dennett (1987): Agents at level 0 decide their behavior without considering behavior of others; Agents at level 1 assume others as level 0 and decide their optimal behavior; Agents at level 2 assume others as level 1 and decide their optimal behavior.

3.5 Experiment

3.5.1 Module Network and Function of Each Module

We constructed module network and the constituent modules based on a FPC model as shown in Fig. 3.5.

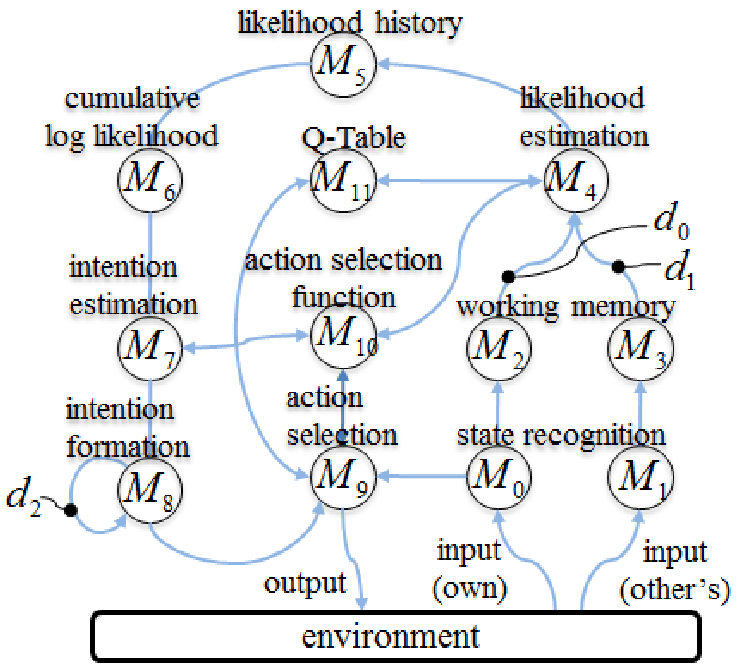


Figure 3.5: The module network used for the experiments.

By combining these modules, above three strategies could be possible (for the sake of simplicity, we defined the link-deactivation signals as described in Fig. 3.5, where $d_0 = d_2$, and $d_0 \neq d_1$). The functions of each module are described as follows.

- $\boldsymbol{M_0}$ (state recognition): Own state $s_s(t)$ and action $a_s(t)$ at the time t are recognized. The state is defined as relative coordinates between self and two preys.
- $\boldsymbol{M_1}$ (state recognition): Other's state $s_o(t)$ and action $a_o(t)$ at the time t are recognized. The state is defined as relative coordinates between others and two preys.

- $\boldsymbol{M_2}$ (working memory): Own state $s_s(t - 1)$ and action $a_s(t - 1)$ at the time $(t - 1)$ are stored.
- $\boldsymbol{M_3}$ (working memory): Other's state $s_o(t - 1)$ and action $a_o(t - 1)$ at the time $(t - 1)$ are stored.
- $\boldsymbol{M_4}$ (likelihood estimation): Likelihood that others' goal (or own goal that would be estimated by others) would be G is estimated as follows:

$$l(G,t) = P(G|s,a) = \frac{\exp(\beta Q(G(t)|s(t-1),a(t-1)))}{\sum_{G'} \exp(\beta Q(G(t)'|s(t-1),a(t-1)))}, \qquad (3.9)$$

  where β is a parameter called the temperature and Q represents an evaluation value which is acquired by reinforcement learning. Likelihood is estimated for all possible goals.

- $\boldsymbol{M_5}$ (likelihood history): In order to make estimation of intention stable, the likelihood $l(G, t)$ is stored in a likelihood history:

$$m_l(G,t) = \{l(G,t), \cdots, l(G,t-H+1)\}, \qquad (3.10)$$

  where H is the length of the history.
- $\boldsymbol{M_6}$ (cumulative log likelihood): Cumulative log likelihood $L(G, t| m)$ (i.e. likelihood that other's or own goal would be G at time t under the condition in which agents have the likelihood history m) is calculated for all possible goals:

$$L(G, t| m) = \Sigma_{l \in m(G,t)} \log l(G, t). \qquad (3.11)$$

- $\boldsymbol{M_7}$ (intention estimation): Others' or own goal G is estimated:

$$g(G,t|m) = \frac{\exp(\beta L(G,t|m))}{\sum_{G'} \exp(\beta L(G',t|m))}. \qquad (3.12)$$

- $\boldsymbol{M_8}$ (intention formation): A hunter judges whether the estimated others' goal G_o and own goal G_s satisfy a cooperated condition. For this task, $G_s \neq G_o$ (i.e. own goal differs from estimated other's goal) is simply a condition of cooperation:
  (1) Own or others' goal is formed to satisfy the cooperated condition.

(2) If the link form M_8 to M_8 is not deactivated, return to (1) just once.

- ***M*** $_{9}$ (action selection): Action a_s is selected:

$$a_s = P(a|s,G)\frac{\exp(\beta Q(a(t)|s(t),G(t)))}{\sum_{a(t)'}\exp(\beta Q(a(t)'|s(t),G(t)))}. \quad (3.13)$$

- ***M*** $_{10}$ (action-selection function): It is the one based on soft-max reinforcement learning.
- ***M*** $_{11}$ (Q-Table): It is an evaluation value acquired by reinforcement learning. Before we conducted experiments, each agent had acquired a different Q-Table on its own by the soft-max reinforcement learning in the setting where there were a hunter and a prey (temperature parameter β =1).

3.5.2 Experimental Setup

We conducted simulations using the hunter task solved by a pair of agents with the module networks. Each agent searches for a set of activation and link-deactivation signals: *Solution* = (*A*, *D*) responding to the current task (or social situation) on the basis of tabu search (Glover, 1997).

First, *N* agents were created, each with a randomly generated initial solution: $Solution_0 = (A_0, D_0)$. Each agent solved the hunter task *N* - 1 times, each with a different agent in a round robin manner. Table 3.1 shows the experimental setup of the experiments. The results obtained were averaged over a time period for each agent.

Table 3.1: Experimental setup.

history size: *H*	5	*T_max*	2
temperature parameter: β	1	*T_min*	-1
$step_{max}$	500	α_e (in searching mode)	0.02
$episode_{max}$	10	α_e (in switching mode)	0.005
$period_{max}$	1000	α_t (in searching mode)	0.02
tabu tenure	7	α_t (in switching mode)	0.005
neighbor size: *n*	13	population size: *N*	8

3.5.3 Results

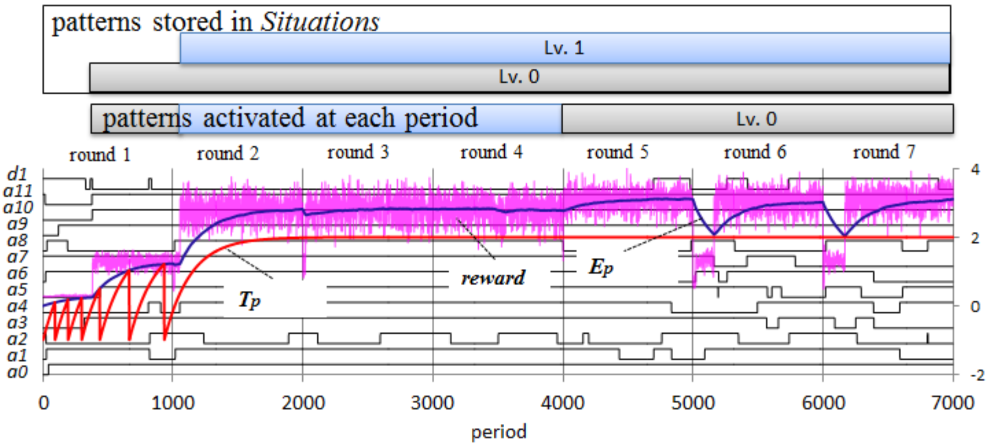
Fig. 3.6-(A) shows the transition of $Solution_p$, *reward*, E_p and T_p of a hunter during learning (7 rounds). In the following account, we call the focal hunter 'Ken'. Fig. 3.6-(B) shows the transition of those of Ken's 7 different partners each round. The each bar above Fig. 3.6-(A) and Fig. 3.6-(B) represent the activated patterns of $Solution_p$ and those in the *Situations*. We have found the emergence and adaptive switching of partial networks for processing higher-order estimation of other's intention in terms of intentional stance developed by Dennett (1987) as shown in Fig. 3.7, which shows the observed networks for processing recursion levels of ToM. In Fig. 3.7, the left networks show the activated patterns of the Ken, whereas the right networks show those of the Ken's partners.

At the 1st period in the 1st round, the *reward* was very low (Fig. 3.6). Fig. 3.7-(A) shows the activated patterns of Ken and his partner at that time. This means that there was no output in the Ken's network while the activated pattern for level 0 ToM was already included in his partner's network. The *reward* increased suddenly at around 400th period, when the activated pattern of the Ken became like Fig. 3.7-(B), in other words, he acquired the network for level 0 ToM. However, the *reward* was not that high because an agent at level 0 continues to chase a specific prey obstinately even if own goal falls on the other's goal.

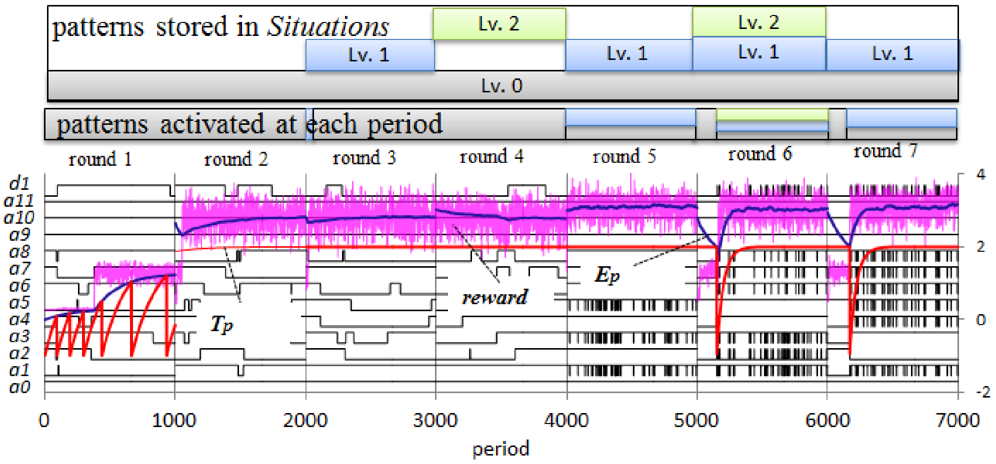
In the 2nd round, Ken and his new partner behaved using level 0 ToM at first as shown in Fig. 3.7-(C). The *reward* increased at around 1050th period in parallel with the activation of Ken's M_4 (Fig. 3.7-(D)). It is because by then Ken had acquired the activated pattern for level 1 ToM.

Between 3th and 5th rounds the *reward* remained stable thanks to the adaptive interactions between the levels of ToM. However, it decreased in the early stage in the 6th round. At that time Ken and his new partner behaved competitively using level 0 ToM. After that, E_p of the partner became less than T_p, and the *reward* increased. This is because the partner shifted into switching mode and reused the adequate activated patterns for the current environment.

Next, we compared the performance with those without the attention system to evaluate the introduction of the mechanism for switching among the kept *Solution*s in *Situations*. Fig. 3.8 shows the transition of the average *reward*s of N agents with and without the attention system, averaged over 10 trials. It is shown that the *rewards* with the system in the early stage of each round were clearly higher than those without it. This indicates that the appropriate partial-network could be swiftly acquired by using the attention system even when interacting a new partner.



(A) A certain hunter (named 'Ken')



(B) Ken's partners which changed per 1000 periods

Figure 3.6: The transition of the Solution$_p$ (black lines), reward (pink line), E_p (blue line) and T_p (red line).

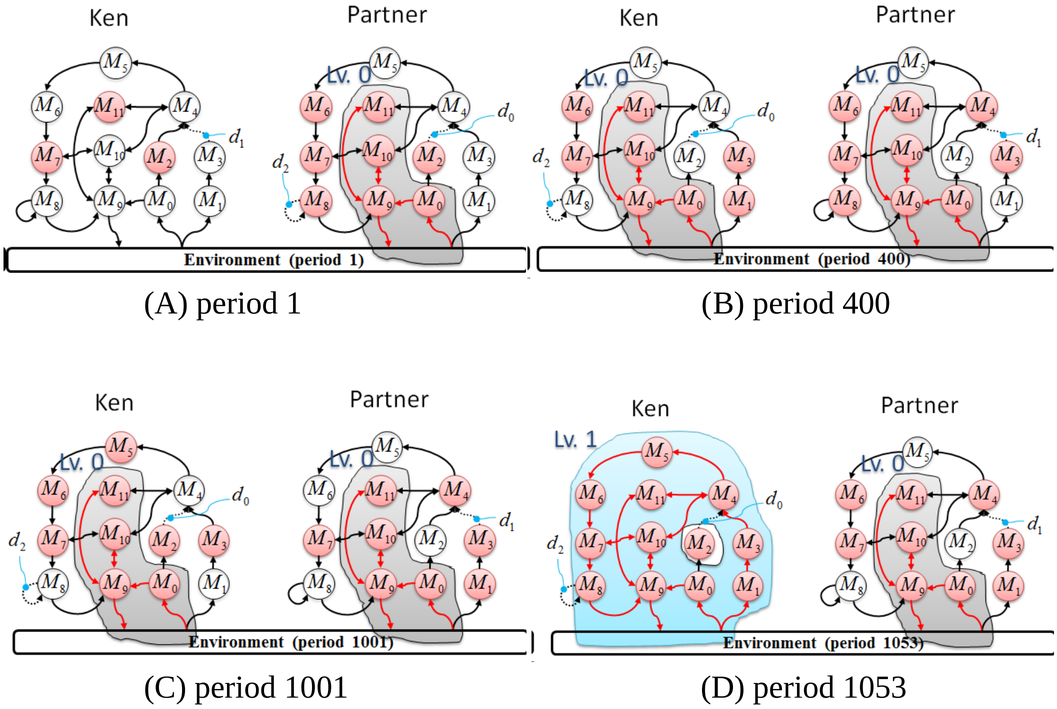


Figure 3.7: The emerged partial networks for processing ToM.

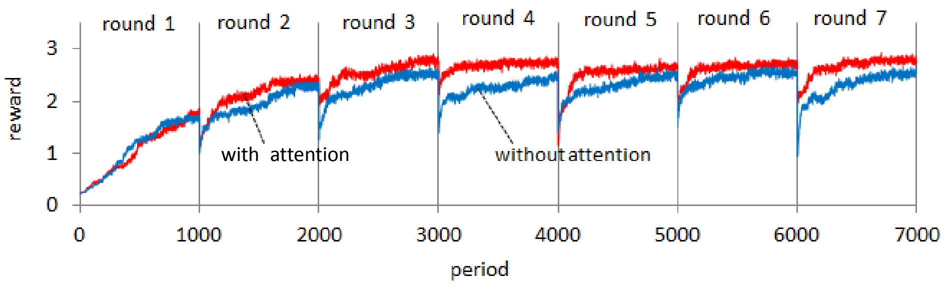


Figure 3.8: The transition of the average rewards of N agents with and without the attention system, averaged over 10 trials.

3.6 Additional Experiment

The result of computer simulation in the previous section showed the acquisition of the effective procedures depending on the interactive agents through learning of the partial networks for processing ToM. In this section, we added some modules and link-deactivation signals into the module network into the previous section in order to make agents behave more flexibly.

3.6.1 Module Network and Function of Each Module

We added some modules for active strategy, called level 2* ToM (Yokoyama and Omori, 2009) into those in the previous section. Agents at level 2* assume others as level 1 and decide their optimal behavior which informs others the own intention explicitly (Fig. 3.9). Explicit action is decided by maximizing the difference in the probability between the own goal G_s and the rest ones $G_{s!}$:

$$a_s = argmax_a(P(G_s|s_s, a_s) - P(G_{s!}|s_s, a_s)). \tag{3.14}$$

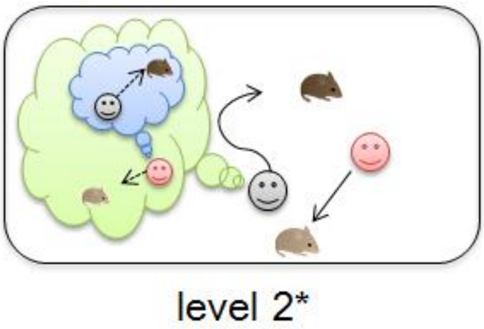


Figure 3.9: Active strategy in which agents assume others as level 1 and decide their optimal behavior which informs others the own intention explicitly.

We constructed module network and the constituent modules based on a FPC model as shown in Fig. 3.10. The functions of each module are described as follows (as to the modules other than M_5 and M_{11}, we reused the same functions in the previous section):

- $\boldsymbol{M_5}$ (likelihood estimation*): Likelihood difference is estimated as follows:

$$d(G,t) = \frac{\exp\left(\beta\left(Q\left(G_i(t)\middle|s(t-1),a(t-1)\right)-Q\left(G_{i!}(t)\middle|s(t-1),a(t-1)\right)\right)\right)}{\sum_{a'}\exp\left(\beta\left(Q\left(G_i(t)\middle|s(t-1),a'(t-1)\right)-Q\left(G_{i!}(t)\middle|s(t-1),a'(t-1)\right)\right)\right)}, \quad (3.15)$$

where $G_{i!}$ represents the goal other than G_i. Likelihood difference is estimated for all possible goals.

- $\boldsymbol{M_{11}}$ (action selection*): Action a_s is explicitly selected by selecting actions that maximize the difference in the probability between the own goal and other than own goal:

$$a_s = \frac{\exp(\beta(Q(G_s(t)|s_s(t),a_s(t))-Q(G_{s!}(t)|s_s(t),a_s(t))))}{\sum_{a'}\exp(\beta(Q(G_s(t)|s_s(t),a'_s(t))-Q(G_{s!}(t)|s_s(t),a'_s(t))))}. \quad (3.16)$$

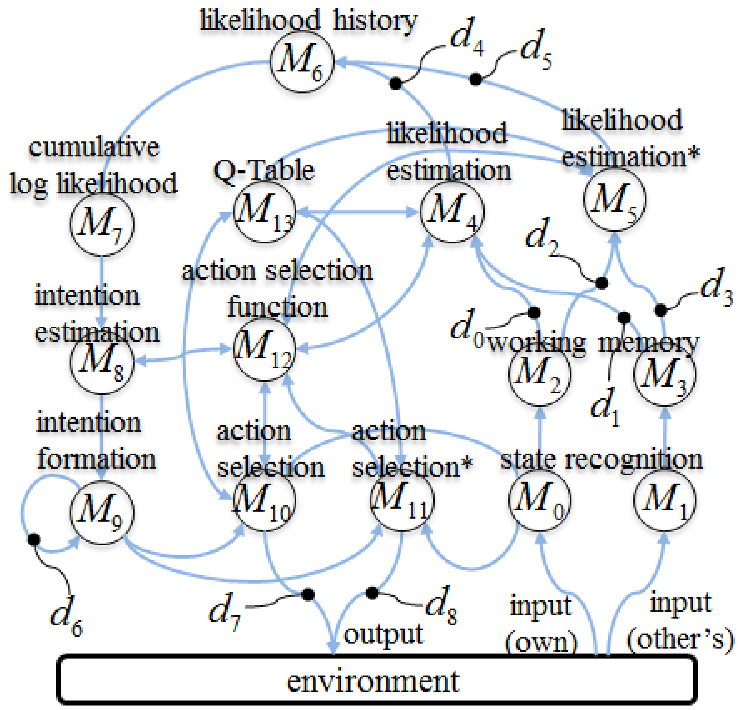


Figure 3.10: The module network used for the experiments.

3.6.2 Experimental Setup

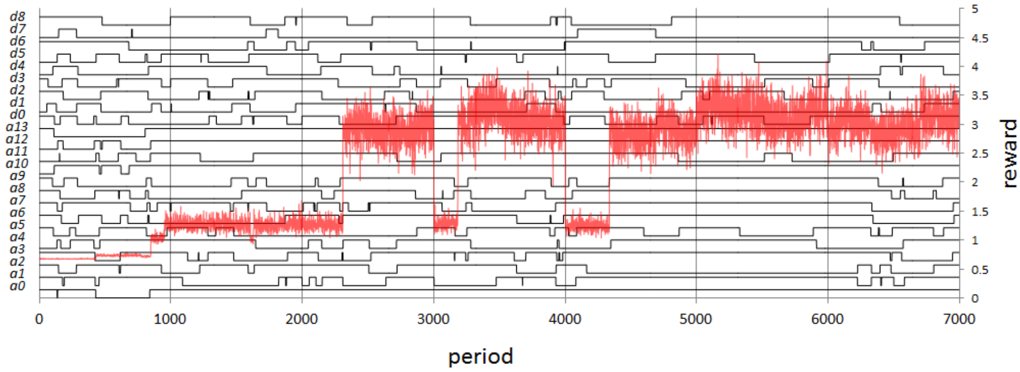
We conducted simulations using the hunter task solved by a pair of agents using the module networks without attention system because there were no differences between the module networks with and without attention system with respect to the basic behavior of the FPC model. The process for searching for a new *Solution* is the same as the previous section. Table 3.2 shows the experimental setup of the experiments.

Table 3.2: Experimental setup.

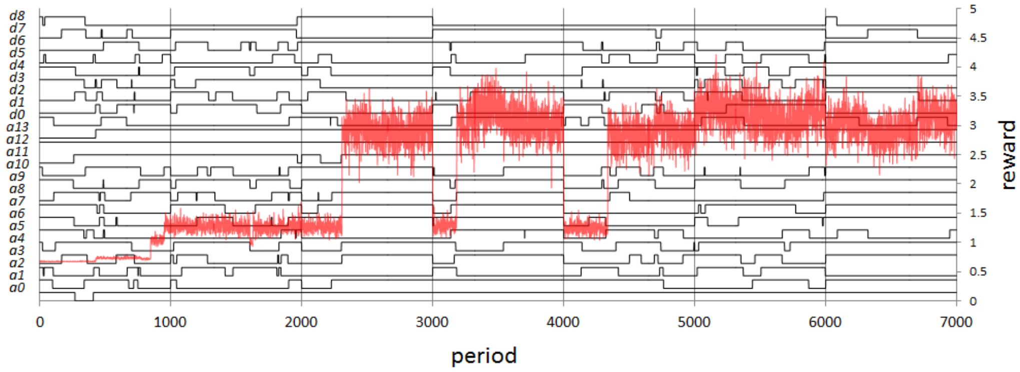
history size: H	5	$episode_{max}$	10
temperature (expression 3.9 and 3.13): β	1/3	$period_{max}$	1000
temperature (expression 3.12): β	2	*tabu tenure*	7
temperature (expression 3.15 and 3.16): β	1	neighbor size: n	23
$step_{max}$	150	population size: N	8

3.6.3 Results

Fig. 3.11-(A) shows the transition of the activation signals, the link-deactivation signal, and the *reward* of a hunter (named Ken) during learning (7 rounds). Fig. 3. 11-(B) shows the transition of those of Ken's 7 different partners each round. In both figures, the red line represents the *reward*, the black lines represent the activation and link-deactivation signals. Fig. 3.12 shows the observed networks for processing recursion levels of ToM (left: Ken, right: partners).

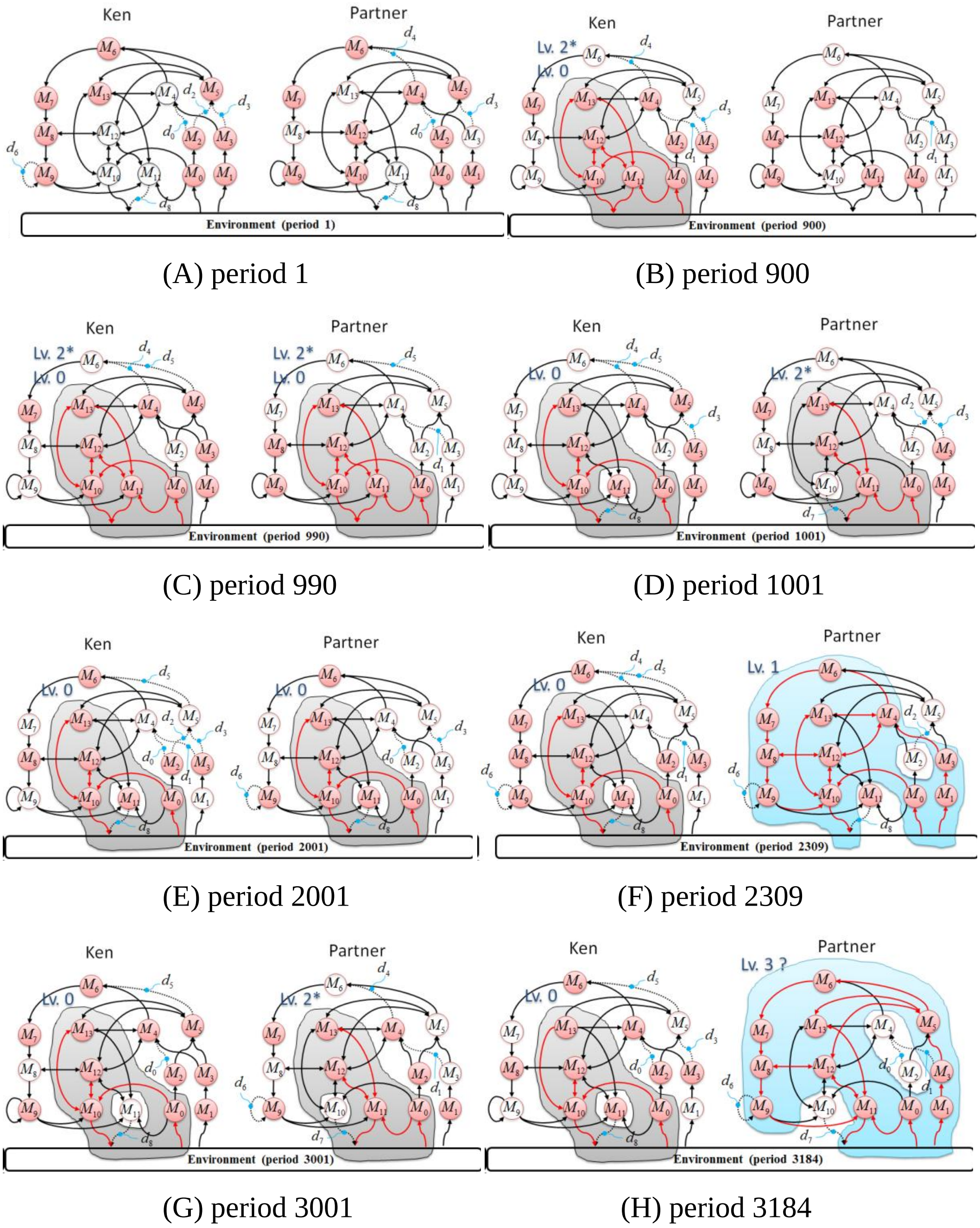


(A) A certain hunter (named 'Ken')



(B) Ken's partners which changed per 1000 periods

Figure 3.11: The transition of the activation signals, the link-deactivation signals, and the reward during the learning (7 rounds).



(A) period 1 (B) period 900

(C) period 990 (D) period 1001

(E) period 2001 (F) period 2309

(G) period 3001 (H) period 3184

Figure 3.12: The emerged partial networks for processing ToM.

At the 1st period in the 1st round, the *reward* was very low because there were no outputs in both hunters' network (Fig. 3. 12-(A)). The *reward* slightly increased at the 900th period in parallel with that Ken acquired the network for level 0 and level 2* ToM (Fig. 3. 12-(B)). Subsequently, the *reward* slightly increased at 990th period. By then, the partner acquired the network for level 0 and level 2* ToM as shown in Fig. 3. 12-(C). At the 1001th period in the 2st round, Ken behaved using level 0 ToM and his new partner behaved level 2* ToM as shown in Fig. 3. 12-(D). The *reward* remained comparatively stable during this round.

At the 2001th period in the 3rd round, Ken and his new partner behaved using level 0 ToM at first (Fig. 3. 12-(E)). After that, there was a remarkable increase in the *reward* at around 2309th period. At that time, Ken had acquired the network for level 0 ToM, whereas his partner acquired the network for level 1 ToM (Fig. 3. 12-(F)). Since then the *reward* remained stable, however; it decreased when the partner changed at the 3001th period because Ken and his new partner behaved competitively using level 0 and level 2* ToM, respectively (Fig. 3. 12-(G)). Then, the *reward* increased at around 3184th period. It is because by then the partner had acquired the network that was functionally similar to level 3 ToM: estimating the goal of others by assuming that others would be at level 2*, and explicitly takes action based on it (Fig. 3. 12-(H)). In the subsequent periods, the *reward* remained stable thanks to the adaptive interactions between the levels of ToM.

3.7 Conclusion and Discussion

We proposed a computational framework for investigating the emergence of ToM as an adaptation on different time scales based on three fundamental hypotheses concerning human brains and cognition. As a first step, this chapter focused on the emergence of ToM to achieve cooperative behavior in a population of agents. The simulation demonstrated a scenario for bootstrapping ToM as the emergence of the partial-networks of functional parts in the brain model based on the interactions between the levels of ToM. It also showed that appropriate behaviors suited for others interacting for the first time can be swiftly acquired simply by reusing the acquired partial-networks. These results imply that efficient social behaviors were attained not only by the individual cognitive components but also the appropriate combinations of these modules, which has much in common with the neuroscientific facts about autism (Happe, 2006). The next step would be to investigate the acquisition of not only the activation signals but also the connections between modules.

It also showed that before an agent acquired activated patterns for level 1 ToM, the partner usually behaved using level 0 ToM. We presume that it (the other behaves using level 0 ToM) is one of the important factors for emergence of the level 1 ToM because it makes agent estimate other’s goal to some extent by using own action selection strategies. This property suggests that the behavior using level 1 ToM further leads to the emergence of level 2 ToM. However, there were only a few cases where a partial network for implementing level 2 ToM emerged in the experiments despite both hunters behaved using level 1 ToM for some time. It might be due to the fact that the number of possible combinations of the activated patterns for level 2 ToM is fewer than those of level 0 ToM. This fact might indicate the difficulty of the emergence of level 2 ToM.

Furthermore, the simulations showed the emergence of the several strategies based on ToM other than the definition of the Dennett's intentional stance such as the extension of active strategy, called level 2* ToM, in which agents lead the intention of others to one's own direction by informing own goal explicitly to others. Fig. 3.13 shows the conceivable definition of the recursion levels of ToM by using the module network adopted in the chapter 3.6. In Fig. 3.13, recursion levels are arranged in ascending order. Rectangular cells represent strategies where agents estimate the intention of others by assuming that others would be at the right overhead strategy. Dark blue rectangular cells mean the possible strategies considering the role division between two agents. Light bule rectangular cells represent those considering cooperation among more than three agents. As shown in Fig. 3.13, the depth of recursion level stops when the agent at an odd level takes action explicitly in the case of cooperation between two agents for the following reason. The explicit action conducted by the agent at the odd level is not transferred to the agent at the next even level (right below in Fig. 3.13) because agents at the even level estimates the intention of others based on own behaviors, independently of those conducted by others. On the other hand, there might be some cases that explicit action taken by the agent at the odd level is effective if there are more than three agents. For example, the agent A (at the odd level *N*) might assume another agent B as even level *N*-1 and the other agent C as odd level *N*-2, and decide own optimal behavior which informs the other agent C the own intention explicitly (Fig. 3.14).

The next step would be to investigate the effectiveness of these active strategies. The situations where agents cannot estimate correctly which preys others aim at supposed to be advantageous to these strategies (Yokoyama and Omori, 2009). These strategies also assumed to be adaptive in the situations where each of two players chooses whether to hunt highly valued prey together and share the proceeds or defect to hunt meager prey of small value, such as a Stag-Hunt game (Yoshida et al., 2008). It might be interesting to discuss the emergence of the cooperative behavior and deception based on the acquisition of the ToM.

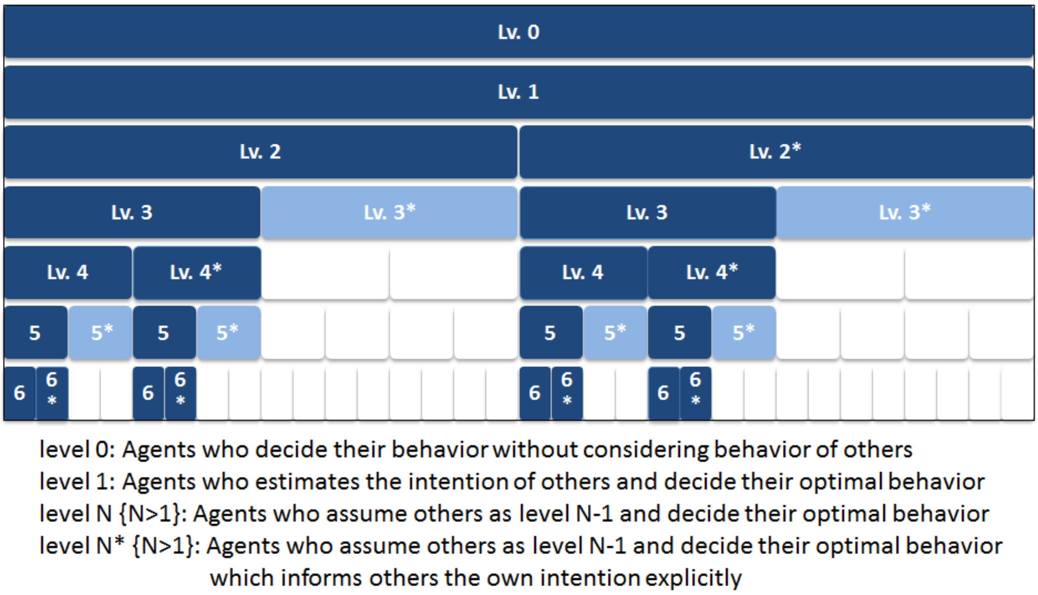


Figure 3.13: Conceivable definition of the recursion levels of ToM by using the module network adopted in the experiments.

We believe that the proposed method would contribute to clarify the origin of ToM. It would also be interesting to discuss the feasibility of the acquisition of ToM in humanoid robots. When applying this model to embodied agents in the real world, there should be a series of challenges they would face, as the functional model of the brain presented in this paper is specialized for the acquisition of ToM for a simple cooperative task in an abstract environment.

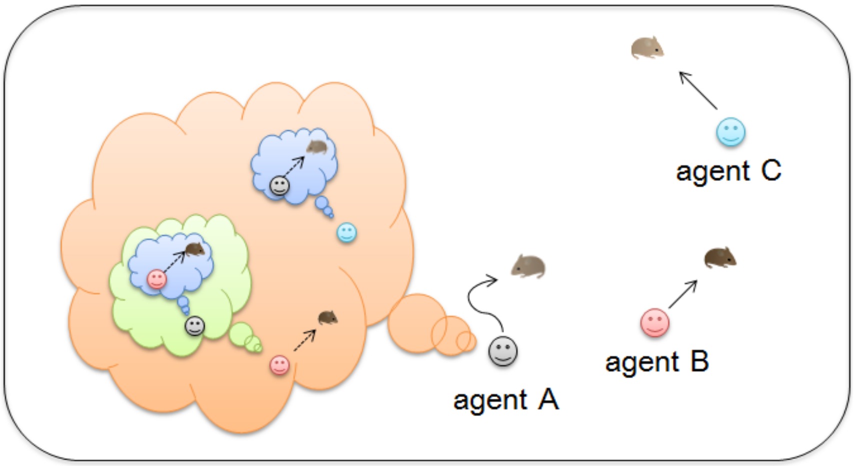


Figure 3.14: Cooperation among three agents.

4 Conclusion

This thesis describes the constructive approach for investigating the emergence of the primitive functions of mind, focusing on the planning for future events and general mind reading ability (Theory of Mind: ToM).

In the second chapter, we discussed the question: what kinds of problems contribute to the emergence of the planning ability? To response to this question, we constructed a blocks world problem as a minimal task which needs ability to think logically, and encoded an inherent planning parameter into the genome. The result of computer simulation showed a general tendency that planning ability emerges when the problem is difficult to solve. When taking social relationships especially in the collective situation into account, planning ability was difficult to evolve in the case that the problem was difficult because there was a conflict between personal and collective interests. Also, the simulation results indicated that sharing information facilitates evolution of the planning ability although the free rider problem tended to be more serious than the situation where agents do not share information. It implies that there is a strong connection between evolution of the planning ability and symbolic communication. It has been claimed that difference in the ecology of the early hominids and the other apes is important but neglected factor in the discussions of the evolution of language (Bickerton, 2002). Our results showed one of the ecologically based answers to why humans are the only animals who have developed a symbolic communication. What remains to be done is to clarify the mechanism of the co-evolutionary dynamics of the planning and symbolic communications in the situation where two kinds of groups, sharing and no sharing information, coexist in the same population.

In the third chapter, we modeled a human communication based on the mind-reading as follows: (1) humans estimate the intention of others from his/her behavior by simulating it based on their own action decision process as if they were in the same situation (Gallese & Goldman, 1998); (2) humans interactively determine own behaviors in accordance with the estimated intention of others. Unlike the static environment where certain rules that agents have to learn were permanent, these processes are too dynamic to be explained by the traditional action learning methods through a trial and error because it is necessary to cope promptly with the change of the intention of others in its dynamic environment. This thesis attempted to demonstrate the emergence of role division in a short time scale in these complex and dynamic social situations. To do this, we used a Functional Parts Combination (FPC) model (Omori and Ogawa, 2001), which regards the brain at a functional level as composed of a set of functional parts and activation signals specifying selectively activated patterns. We conducted computer simulations in which the activation signals were learned using a hunter task as a problem to be solved by the agents. The simulation demonstrated a scenario for bootstrapping ToM as the emergence of the partial-networks of functional parts in the brain based on the interactions between the recursive levels of intentionality in ToM. It also showed that appropriate behaviors suited for others interacting for the first time could be swiftly acquired simply by reusing the acquired partial-networks. The next step would be to consider the more complex environment: real world. When applying this model to embodied agents in the real world, there should be a series of challenges they would face. One important problem that comes to mind is a notorious *symbol grounding problem* (Harnard, 1990), i.e., how symbols get their meaning, or how the connection between symbols and the environment can be established. Another problem is that agents cannot adapt various situations. To solve this problem, it would be necessary both to add the functional parts responding to the several situations and to increase the particle size of themselves for flexible behavior. However, it might generate a new problem: (1) it is difficult to understand how complex structures and behaviors emerge from them; (2) how efficiently combinations of the lots of modules search, in other words; *frame problem* (McCarthy and Hayes, 1969).

Acknowledgement

I would like to give heartful thanks to Prof. Arita whose comments and suggestions were innumerably valuable throughout the course of my study. I would also like to thank Prof. Suzuki, Prof. Unemi, Prof. Omori and Prof. Kita who gave me invaluable comments and warm encouragements. I also appreciate the feedback offered by all the people in the lab. I would also like to gratefully express my gratitude to my family and Soka University Founder Daisaku Ikeda for their moral support and warm encouragements. Finally, I appreciate the financial support of Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology (No. 21120010), and by the Sasakawa Scientific Research Grant from The Japan Science Society that made it possible to complete my thesis.

References

Alexander, R. D. (1971). The search for an evolutionary philosophy of man. *Proceedings of the Royal Society of Victoria*, 84, 99–120.

Alexander, R. D. (1990). *How did humans evolve? : Reflections on the uniquely unique species.* Ann Arbor, Michigan, USA: University of Michigan Museum of Zoology (Special Publication No. 1).

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, 4 (7), 267-278.

Anderson, J. R. (1980). *Cognitive Psychology and its Implications*. W. H. Freeman and Company.

Arita T., Unno K., & Kawaguchi K. (1995). A primitive model for language generation by evolution and learning. *International Workshop on Biologically Inspired Evolutionary Systems*, 181-186.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4 (11), 417-423.

Bickerton, D. (2002). Foraging versus social intelligence in the evolution of protolanguage. In A. Wray (Ed.), *The Transition to Language* (pp. 207-225). Oxford: Oxford University Press.

Bisig, D., & Pfeifer, R. (2008). Understanding by design: the synthetic approach to intelligence. In: *Explorations in architecture: teaching, design, research*. Birkhauser, Basel.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulated cortex: an update. *Trends in Cognitive Sciences*, 8 (12), 539–546.

Braitenberg, V. (1984). *Vehicles: experiments in synthetic psychology*. MIT Press.

Brinck, I., & Gärdenfors, P. (2003). Co-operation and communication in apes and humans. *Mind and Language*, 18, 484-501.

Brothers, L. (1990). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, 1, 27–51.

Brune, M., & Brune-Cohrs, U. (2006). Theory of mind – evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience and Behavioural Reviews*, 30, 437-355.

Byrne R. W. (1997). The technical intelligence hypothesis: an additional evolutionary stimulus to intelligence? In: Whiten A., & Byrne R. W. (Eds.), *Machiavellian intelligence. vol. II: extensions and evaluations* (pp. 289-211). Cambridge : Cambridge University Press.

Brodmann, K. (1925). *Vergleichende Lokalisationslehre der Grosshirnrinde*. Leipzig: Barth.

Carruthers, P. (1996). Theories of theories of mind. In: A *defence of the theory-theory (chapter Simulation and self-knowledge)*. Cambridge: Cambridge University Press.

Clark, D. D., & Sokoloff, L. (1999). in *Basic Neurochemistry: Molecular, Cellular and Medical Aspects*, eds. Siegel, G. J., Agranoff, B. W., Albers, R. W., Fisher, S. K., & Uhler, M. D. (Lippincott, Philadelphia), 637–670.

Craik, K. (1943). *The nature of explanation*. Cambridge: Cambridge University

Press.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7 (9), 415-423.

Dennett, D. (1987). *The Intentional Stance*. Cambridge: MIT Press.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.

Deacon, T. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: W. W. Norton.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6 (5), 178–190.

Evans, D., & y de Back, W. (2003). *Synthetic evolutionary psychology*. Consultado: 17 junio 2007, de Dylan Evans' Home Page. Sitio web: http://www.dylan.org.uk/syntheticEP14.html.

Flinn, M. V., Geary, D. C., & Ward, C. V. (2005). Ecological dominance, social competition, and coalitionary arms races: why humans evolved extraordinary intelligence. *Evolution and Human Behavior*, 26, 10-46.

Gallese, V., & Goldman A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2, 493–501.

Gärdenfors, P. (2003). *How homo became sapiens: on the evolution of thinking*. Oxford: Oxford University Press.

Geary, D. C. (2004). *The origin of the mind: evolution of brain, cognition, and general intelligence*. American Psychological Association.

Glover, F., & Laguna M. (1997). *Tabu Search*. Norwell: Kluwer Academic Publishers.

Gordon, R. M. (1996). Sympathy, simulation, and the impartial spectator. *Ethics*, 105, 727–742; reprinted in May L., Friedman M., & Clark A. (Eds.), *Mind and Morals: Essays on Ethics and Cognitive Science*. Cambridge: MIT Press.

Gulz, A. (1991). The planning of action as a cognitive and biological phenomenon. Lund: *Lund University Cognitive Studies 2*.

Happe, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience*, 9, 1218-1220.

Harbers, M., Bosch, K. van den & Meyer, J-J. Ch. (2009). Modeling agents with a theory of mind. *Proeedings of International Conference on Intelligent Agent Technology*, 217-224.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162 (3859), 1243-1248.

Harnad, S. (1990). The Symbol Grounding Problem. *Physica D*, 42, 335-346.

Hill, K. (1982). Hunting and human evolution. *Journal of Human Evolution*, 11, 521–544.

Hirasawa, K., Okubo, M., Hu, J., & Murata, J. (2001). Comparison between Genetic Network Programming (GNP) and Genetic Programming (GP). *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, 1276-1282.

Humphrey, N. K. (1976). The social function of intellect. In: P. P. G. Bateson, & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge: Cambridge University Press.

Jolly, A. (1999). *Lucy's legacy: sex and intelligence in human evolution.* Cambridge: Harvard University Press.

Kirkpatrick S., Gelatt C. D., & Vecchi M. P. (1983). Optimization by simulated annealing. *Science*, 220 (4598), 671–680.

Koza, J. R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge: MIT Press.

Langton, C. G. (1992). Artificial life. In: Nadel L. & Stein D. L. (Eds.), *Lectures in complex systems*. Reading: Addison-Wesley.

Levy, R., & Goldman-Rakic, P. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. *Experimental Brain Research*, 133, 23–32.

Langton C. G. (1987). Artificial life. *The proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems*, Addison-Wesley Publishing Company, Inc.

Liu, M. J., Fenwick, P. B. C., Lumsden, J., Lever, C., Stephan, K.-M., & Ioannides, A. A. (1996). Averaged and single-trial analysis of cortical activation sequences in movement preparation, initiation, and inhibition. *Human Brain Mapping*, *4*, 254–264.

McCarthy J., & Hayes P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence*, 4, 463-502.

Miwa K. (1999). Kagakuteki hakken ni okeru yuukouna kasetukensyou houryaku: keisanki simure-syon ni motozuku kentou (in Japanese). Okada T., Todayama K., Tamura H., & Miwa K. (Hen) : *Kagaku wo kangaeru: Jinkou chinou kara karutyuraru sutadi-zu made no 14 no siten*, kita ooji syuppan.

Moreno, A. (2002). Artificial Life and Philosophy. *Leonardo*, 35(4), 401-405.

Nagata, Y., Ishikawa, S., Omori, T., & Morikawa, K. (2010). Computational model of cooperative behavior based on dynamical selection of intention based action decision strategy (in Japanese). *Cognitive studies*, 17(2), 280-286, 2010.

Namatame A. (Hen) (2004). *Ge-muriron to sinkadainamikusu* (in Japanese). Tokyo: Morikita Syuppan.

Ney, H., Mergel, D., Noll, A., & Paeseler, A. (1992). Data driven organization of the dynamic programming beam search for continuous speech recognition. *IEEE Transsactions of Signal Processing,* SP-40 (2), 272-281.

Noble, J., Hebbron, T., Horst, J., Mills, R., Powers, S., & Watson, R. (2010). Selection pressures for a Theory-of-Mind faculty in artificial agents. *Proceedings of the 12th International Conference on the Simulation and Synthesis of Living Systems (ALIFE XII).*

Ogawa, A., & Omori, T. (2002). Looking for a suitable strategy for each problem - Multiple tasks approach to navigation learning task. *Proceedings of the 2nd International Workshop on Epigenetic Robotics*, 125 - 132.

Ogawa K., Inui T., & Sugio T. (2006). Separating brain regions involved in internally guided and visual feedback control of moving effectors: an event-related fMRI study. *Neuroimage*, 32, 1760-1770.

Oliveri, M., Turriziani, P., Carlesimo, G., Koch, G., Tomaiuolo, F., Panella, M., & Caltagirone, C. (2001). Parieto-frontal interactions in visualobject and visual-spatial working memory: evidence from transcranial magnetic stimulation. *Cerebral Cortex*, 11, 606–618.

Omori, T., & Ogawa, A. (2001). Two hypothesis for realization of symbolic processing in brain. *Proceedings of the 9th International Conference on Neural Information Processing*.

Osvath, M., & Gärdenfors, P. (2005). Oldwan culture and the evolution of anticipatory cognition. *Lund University Cognitive Science*, 122.

Pfeifer R., & Bongard J. C. (2007). *How the Body Shapes the Way We Think—A New View of Intelligence*. Cambridge: MIT Press.

Potts, R. (1998). Variability selection in hominid evolution. *Evolutionary Anthropology*, 7, 81–96.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, 1, 515–523.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.

Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O., & Raichle, M. E. (1992). Activation of the hippocampus in normalhumans: a functional anatomical study of memory, *Proceedings of the National Academy of Sciences of the United States of America*, 89, 1837–1841.

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of human mind. *Genetic, Social and General Psychology Monographs*, 123, 133-167.

Takahashi, H., Ishikawa, S., & Omori, T. (2008). Modeling of dynamical cognitive control mechanism that supports social interaction (in Japanese). *Cognitive Studies*, 15(1), 202-215.

Takano, M., & Arita, T. (2006). Asymmetry between even and odd levels of recursion in a Theory of Mind. *Proceedings of the 10th International Conference on the Simulation and Synthesis of Living Systems (ALIFE X)*, 405-411.

Toga, A. W., &Maziotta, J.C. (2000). *Brain Mapping - The systems*. San Diego: Academic Press.

Tooby, J., & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modelling. In Kinzey, W (Ed.), *The evolution of human behavior: Primate models* (pp. 183-238). New York: State University of New York Press.

Trealeven, P. C., Hopkins, R. P., & Rautenbach, P. W. (1982). Combining data flow and cotrol flow computing. *Computer Journal*, 25 (2), 207-217.

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, 2, 67-70.

Wolpert, D. M., Goodbody, S. J., & Husain, M. (1998). Maintaining internal representations: The role of the human superior parietal lobe. *Nature Neuroscience*, 1 (6), 529-533.

Yokoyama, A., & Omori, T. (2009). Model based analysis of action decision process in collaborative task based on intention estimation (in Japanese). *The Transactions of the Institute of Electronics, Information and Communication Engineers. A*, J92-A (11), 734-742.

List of Publications

- Journal

  Minoya, K., Arita, T., & Omori, T. (2011). Autonomous acquisition of cooperative behavior based on a theory of mind using parallel Genetic Network Programming. *Artificial Life and Robotics*, 16 (2), 157-161.

  Minoya, K., Unemi, T., Suzuki, R., & Arita, T. (2011). A constructive approach to the evolution of the planning ability. *International Journal of Artificial Life Research*, 2 (3), 22-35.

- Refereed International Conference Proceedings

  Minoya, K., Arita, T., & Omori, T. (in press). A constructive approach for investigating the emergence of role division in social interactions. *Advances in Cognitive Neurodynamics* (*III*) (*The 3rd International Conference on Cognitive Neurodynamics*).

  Minoya, K., Arita, T., & Omori, T. (2011). An artificial life approach for investigating the emergence of a Theory of Mind based on a functional model of the brain. *Proceedings of the 2011 IEEE Symposium on Artificial Life*, 108-115.

  Minoya, K., Arita, T., & Omori, T. (2011). Autonomous acquisition of cooperative behavior based on a theory of mind using parallel Genetic Network Programming. *Proceedings of the 16th International Symposium on Artificial Life and Robotics*, 456-459.

  Minoya, K., Unemi, T., Suzuki, R., & Arita, T. (2009). A constructive approach to the evolution of the planning ability. *Proceedings of 13th Asia Pacific Symposium on Intelligent and Evolutionary Systems*, 38-44.

- National Congress

Omori, T., Minoya, K., & Arita, T. (2011). Brain circuit search and symbol emergence for flexible action decision (in Japanese) *The 25th Annual Conference of the Japanese Society for Artificial Intelligence*.

Minoya, K., & Unemi, T. (2009). A fundamental study on evolutionary acquisition of planning ability (in Japanese). *Proceedings of 36th Intelligence System Symposium*, 36, 79-84.

Minoya, K., & Unemi, T. (2008). Pulan-ning no kigen ni tsuite (in Japanese). *14th emergent system symposium, summer school of the emergent system*.

Minoya, K., & Unemi, T. (2008). Pulan-ning no kigen ni tsuite (in Japanese). *Meeting for Youth Community (MYCOM 2008).*