A Constructive Approach to the Evolution of the Planning Ability

Kenichi Minoya 1, Tatsuo Unemi 2, Reiji Suzuki 3, and Takaya Arita 4

1,3,4 Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
Email: 1 kenichiminoya@alife.cs.is.nagoya-u.ac.jp
3,4{reiji, arita}@nagoya-u.jp
2 Department of Information Systems Science, Soka University
Tangi-machi 1-236, Hachioji, Tokyo 192-8577, Japan
Email: unemi@iss.soka.ac.jp

Abstract: This paper describes the first stage of our study on evolution of the planning abilities. We use a blocks world problem as the task of agents, and encode the inherent planning parameters into the genome. The result of computer simulation indicated that planning ability emerges when the problem becomes difficult to solve. Moreover, the result clarified that the balance between the costs of action and thinking is important for the emergence of planning abilities. Furthermore, it has been implied that there might be a strong connection between the evolution of symbolic communication and planning ability in terms of collective situations.

Keywords: planning, evolution, ecological pressure, social pressure, symbolic communication.

1. Introduction

Humans have lots of characteristics that distinguish us from other species. One of the most distinctive features is the ability to make plans for future. Meanwhile there are several clear cases of planning among nonhuman primates [1]. For example, chimpanzees perform the following sequence of actions: walk away from a termite hill, break a twig, peel this leaves off to make a stick, return to the termite hill, and use the stick to "fish" for termites [2]. However, tool making behavior of these primates occurred only in the presence of a visible reward, and never without it [2]. According to the Bischof-Kohler hypothesis, humans are the only extant animals that can plan for future needs [3]. Gulz (1991, p. 55) calls planning for present needs immediate planning while planning for future needs is called anticipatory planning [4]. So why might humans, but not other animals, have the abilities to make plans for future needs?

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 pressures drove the evolution of intelligence of human. For example, species which confront complicated feeding and predation pressure have larger content of brain and higher EQ than those of simple feeding and predation pressure [5]. Also, Gardenfors (2005) argued that Oldowan cultural niche that is based on transport over extended space and time promoted the selection for anticipatory planning [2].

Ecological pressures produce abilities to get a resource such as prey from the environment and to prevent the use from a predator as a resource [6] (p. 232). Then, the question is: what kinds of environment have generated the abilities for anticipatory planning in the hominid line? In the sphere of the Artificial Intelligence, environment can be defined by the state space. Fig. 1 shows part of the state space of the Blocks World Problem (see Section 2.2 for detail). To investigate the ecological pressures which contribute to the emergence of planning abilities, we focus on the properties of the state space which decide the complexity of a given problem.

Fig. 1 Part of the state space of blocks world problem.

As the problem becomes more complex, action cost might be the most critical factor to survive because individuals who take lots of actions in vein consumes extra energy and will take long time to achieve a objective such as hunting any animals for food. However, we cannot ignore the cost of planning because big brains are exceedingly expensive both to evolve and maintain. We assume influence of cost of planning has been changed in the hominids line in several contexts. For example, changing a diet from predominantly vegetarian to more protein and fat based would supply the
nutrients to the brain, and then it might exceed the planning cost. Also, if we had a capacity to engage in a cognitive task quickly, correctly, and economically, the influence of cost of planning would be weak. For these reasons, we focused on the cost of action and planning which act on individual survival, and expressed the balance between the costs of action and planning, as shown in Fig. 2. In this context, as the influence of the cost of planning decreases, the influence of the cost of action increases, and vice versa.

![Fig. 2 The balance between the cost of action and planning.](image)

Yet, common problems for these ecological models include difficulties in explaining why humans evolved such extraordinary cognitive competencies, considering that many other species hunt, occupy savanna habitats, endured the same climatic fluctuations, and so forth [7]. 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 [8] (p. 4-7). Co-operating with other people is considered to be one of the most important factors to deal with fellow in social circumstances. Furthermore, symbolic communication seems to be indispensable to co-operate smoothly with other individuals. Gärdenfors et al. (2003) traced the difference between the ways in which apes and humans co-operate to differences in communicative abilities, claiming that there is a strong connection between the evolution of anticipatory planning and symbolic communication [9]. However, there is little known about the specific mechanisms that underlie it.

In this study, we construct a blocks world problem consisting of a set of rectangular blocks, a table, and some agents, and explore the dynamics in the mechanism of evolution of the anticipatory planning. Our first goal is to elucidate the environment which drove the evolution of planning ability. Our second goal is to investigate the relation between the evolution of planning ability and symbolic communication. Section 2 explains the planner, the task, and the architecture of agents. Section 3 shows the experiments about the ecological pressure, and section 4 describes the experiments about the social pressure. Section 5 summarizes the paper.

2. The Model

2.1 Planner

Gärdenfors (1995) distinguished between two kinds of representations of information: A cued representation stands for something that is present in the current situation; a detached representation stands for objects or events that are neither present in the situation nor triggered by some recent situation [10]. Also, immediate planning only requires a cued representation of the current needs, however; anticipatory planning requires a detached representation of its future needs [10]. In the following model, we define the inherent planning parameter as an attribute value which enables anticipatory planning.

A beam search algorithm [11] is used for the planner of each agent (see Fig. 3). Beam width is defined as 7 in the following experiments. In the process of beam search, if the number of times of expanding is greater than an inherent planning parameter (we call "planning limit" hereafter), planner returns the solution by tracing the path from the state which has the best value in the priority queue (as a short term memory) to the present state. Therefore, the more the number of times of expanding is, the deeper the search is. For example, in case the planner expands only 1 time, it generates S → A or B. Also, in case planner expands 4 times, it generates S → B → D → H → I. In other words, long planning limit enables agents to take detached representations for future goals.

![Fig. 3 The beam search adopted in the model. Alphabets represent states. Among them, S represents the start state. Figures represent heuristic values representing the approximate steps from the focal state to the goal state, by which partial solutions are evaluated. The states in yellow squares are those stored in the priority queue. If the number of stored states is greater than the beam width (4 in this example), the node with the worst score is removed from the queue, and the best node in the queue is expanded and then removed. The successor nodes are added to the queue if they are not there. In this manner, the best node in the current queue is expanded repeatedly (if all nodes have same heuristic values, the planer randomly expands next states) until the number of times of expanding is greater than the inherent attribute value of the agent ("planning limit") and then the solution is obtained by tracing the path from the best node in the queue to the start node.](image)

2.2 Task and Architecture

We conducted simulations using a classical blocks world problem as the task assigned to agents. A blocks world consists of a limited space of table with five rectangular blocks and a grip. An agent is allowed to move a block to the top of another block stack 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, in case the table size is large, agents have many choices when they move a block. Given the current and target configurations of the blocks, the blocks world problem asks for a sequence of grip motions to achieve the target configuration.

The architecture is shown in Fig. 4. First, an agent figures out the present state, and passes it to the planner. Then, the planner makes a plan by applying beam search. If the number of times of expanding reaches the planning limit, the planner stops searching, and returns the solution. Next, the agent moves blocks by the gripper. We define an action step as a movement of a block by a gripper. Action steps are repeated until the target configuration is achieved or they exceed the upper limit (500).

![Fig. 4 Architecture of each agent.](image)

3. Ecological Pressure and Planning

3.1 Basic Experiments

We conducted basic experiments to find how obtained solution is influenced by the difference in planning limits and the difficulty of a given problem. The difficulty of a problem was changed by the two properties: the depth and the size of the table. The depth is defined as the shortest path from start to goal. We created 4 settings where the depth is fixed (depth 18), but with a varying table size 4, 5, 6, and 7 (start and goal configuration is fixed). Also, we created 4 settings where the size of the table is fixed (table size 7), but with a varying depth 8, 12, 16, and 18 by changing the start configuration (goal configuration is fixed). The following results have been averaged over 100 trials.

Fig. 5 shows the average action steps to solve a problem in various settings of the table size, the depth of the problem, and the planning limit. As shown in Fig. 5, solution became depleted when the table size is small and depth is long in a case in which the planning limit is short. This is because problem became more difficult, and then it is necessary to make plan deeply in those situations.

![Fig. 5 Average action steps to solve a problem in various settings of the table size, the depth of the problem, and the planning limit.](image)

3.2 Evolutionary Experiments

We conducted simulations in which the planning limit of agents was evolved by using a genetic algorithm. A chromosome was represented by integer encoding. We first created 20 individuals whose planning limits were randomly selected from 1 to 3.

Each agent solved blocks world problem 10 times, and was evaluated by the fitness function:

$$F = w_a \cdot \frac{a}{a} + w_p \cdot \frac{p}{p} \quad \left( w_a + w_p = 1 \right),$$

where $a$ was average action steps of 10 trials of each agent, $p$ was planning limits of each agent, the $\bar{a}$ and the $\bar{p}$ was a fixed number ($\bar{a} = 100$, and $\bar{p} = 10$), $w_a$ was the weight to the action cost, and $w_p$ was the weight to the planning cost.

Expression (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 was determined by $w_p$ and $w_a$.

The offsprings in the next generation were selected by the ranking selection. In the ranking selection, each rank had the selection probability, and each individual was ranked by scaled adaptive value $f'_i$:

$$f'_i = (N - R_i + 1)^2,$$

and selection probability $p_i$ was defined as:

$$p_i = \frac{f'_i}{\sum_{j=1}^{N} f'_j},$$

where $N$ was the population size and $R_i$ was the rank of individual $i$. Then, the gene of all offsprings was mutated with a probability 0.5. In the phase of mutation, a random integer digit $m$ (uniform distribution from $-8$ to $+8$) was generated, and added to the original gene. If the result was less than 1, the gene became $1 - m$.

We conducted evolutionary experiments where the depth had a fixed parameter (depth 18) but with a varying size of table 4, 5, 6, and 7. We also conducted experiments where the table size had a fixed parameter (table size 7) but with a varying depth 8, 12, 16, and 18. Besides, the balance between the cost of action and planning $(w_a, w_p)$ was changed in all 8 environments. The following result is the average of 10 trials for 500 generations.
3.3 Results

Fig. 6 and Fig. 7 show the difference in transition of the fitness, action steps, and planning limits of the best individual (who has the best fitness among individuals) in various settings of the table size (Fig. 6) and the depth of the problem (Fig. 7) when the balance between the cost of action and planning ($w_a$ and $w_p$ in the expression (1)) was changed. The basic finding is that action steps decreased and planning limit increased through the course of evolution. However, as the weight of the planning cost increased, planning ability did not emerge when the table size was large (bottom right in Fig. 6 is an example where $w_p$ was 0.16). This is because large size of the table and high cost for planning negated beneficial effect on the planning. Also, we found that the longer the depth from start to goal was, the longer the planning limit was (top right in Fig. 7 is an example where $w_p$ was 0.04). However, as the weight of the planning cost increased, planning ability did not emerge when the depth from start to goal was long (bottom right in Fig. 7 is an example where $w_p$ was 0.20). This is because agents in the shallow environment can minimize action steps by the short planning limit, however; agents in the deep environment need longer planning limit to minimize action steps. It implies deep environment drove the evolution of the planning, however; planning cost was strongly worked in those environments.

![Fig. 6 Difference in transition of the best fitness, action steps, and planning limit between small and large sizes of the tables when the weight of the planning cost ($w_p$) was changed. Top row: $w_p=0.04$; bottom row: $w_p=0.16$.](image)

![Fig. 7 Difference in transition of the best fitness, action steps, and planning limit between shallow and deep environment when the weight of the planning cost ($w_p$) was changed. Top row: $w_p=0.04$; bottom row: $w_p=0.20$.](image)
4. Social Pressure and Planning

When problems are difficult such as making complicated stone tools, making a tent, farming, or stock raising, working together seems to be important to accomplish a goal efficiently. Furthermore, it might be difficult to collectively work without an advanced form of communication in those situations. Therefore, increasing symbolic communication skills might be favored because such skills allowed individuals to better anticipate and influence social interactions with other increasingly intelligent individuals. The main goal of this section is to investigate the relation between the evolution of planning and symbolic communication.

4.1 Model

We constructed two models, agents in the first one work together to reach the same goal, but do not share information, and agents in the second one also work together to reach the same goal by sharing information. We call the former the nonshared model, and the later the shared model. In both models, we assumed a situation in which two agents participate in a collective task. We call the first one agent “A”, and the second one agent “B”. In the nonshared model, which is like a turn-taking game, the agent A makes a plan by beam search until the number of times of expanding reaches the inherent planning limit. Then he moves blocks together with agent B. This cycle is repeated until action step reaches the determined number of times (7 times). When action steps reaches this number, the agent B changes places with the partner A, makes plan, and moves blocks together with agent A. This cycle is repeated until action step reaches the same of times. This phase is repeated until the agents accomplish a goal, or exceeds an upper limit of action step (500 steps).

On the other hand, in the shared model, both agents make plan at the same time, and compare their plans. In a case in which the heuristic values of both agents are the same, the plan which has shorter sequence of actions is selected. In addition, the length of both sequences are the same, either one is randomly selected. When an appropriate plan is selected by mutual agreement, both agents move blocks together according to the plan. This phase is repeated until the agents accomplish a goal, or exceeds an upper limit of action step (500 steps). This model assumes symbolic communication is used when they compare plans.

4.2 Basic Experiments

We conducted basic experiments to find how obtained solution is influenced by the difference in how to co-operate. The depth and table size of the environment in both models were 18 and 7, respectively. Fig. 8 shows the action steps averaged over 100 trials of the task evaluation when varying the planning limits of two agents (left: nonshared model; right: shared model). The x and y axes correspond to the planning limit of agent A and that of agent B, and the z axis (color) represents the average action steps. 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 short.

Also, comparing the nonshared model with the shared model, agents in the shared model could decrease action steps by the shorter planning limit than those in the nonshared model for the following reason. At the process of the planning, the planner randomly selected actions in a case in which the heuristic values in the priority queue are the same, and it varied in plans. Because a better of them is selected, agents in the shared model could behave more efficiently by comparing both plans.

4.3 Evolutionary Experiments

We conducted simulation in which the planning limit of agents was evolved by using a genetic algorithm. We simulated an environment where the depth from start to goal was 18, and the table size was 7. Besides, in both experiments, \( w_p \) was changed, ranging from 0.0 to 0.20 with 0.02 increments in between. As the planning architecture, we used the same fitness function and genetic operations as those used in the experiments on ecological pressure. Every pair of agents solved the problem in a round robin manner, and action steps were averaged over those games in both models. The following result is the average of 10 trials for 500 generations.

4.4 Results

Fig. 9 shows the differences in transition of averages of the fitness, action steps, and planning limit between basic model (where agents solved tasks alone), nonshared model, and shared model when the weight of planning cost (\( w_p \)) was changed. It is shown that action steps decreased, and planning limit increased through the course of evolution, however; an increase in the weight of the planning cost made the planning limit of agents shorter in the nonshared model (bottom right in Fig. 9 where \( w_p = 0.12 \)). This is because there was necessity to take longer planning limit to minimize action steps in the nonshared model, however; high cost of planning made the planning limit of agents shorter. Also, 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. 8). Therefore, in a case in which agents who had long planning limit occupied a major part of population, agents who had extremely 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 [12] which explains the reason why the planning ability did not emerge in the collective situation.

The notable point is that planning ability emerged in shared model although the free rider problem tends to be more serious when there was a certain degree of the weight of the planning cost (bottom right in Fig. 9). This is because agents could minimize action steps with the small planning limit because sharing information improved a quality of the solution. Fig. 10 shows the transition of the distribution of the planning limit among individuals of a certain trial in shared model when \( w_p \) was 0.12. There were two general phases in dynamics as follows. First, agents who had extremely short and long planning limit coexisted at the same generations (phase 1). Next, agents who had extremely short planning limit disappeared, and agents who had a long planning limit almost-totally occupied among population (phase 2). This phenomenon could be detected in all 10 trials when there was a certain degree of the weight of the planning cost (\( w_p \) is in the range of 0.12 to 0.14) in the shared model. This result might indicate the extinction of several species of Homo as a result of the social competition within groups.

5. Conclusion

Simulation results on ecological pressure and planning indicated that planning ability emerges in the environment where the table size is small and the optimal path of the solution from start to goal is long. This is because a problem becomes hard to solve in those two situations. However, an increase in the weight of the planning cost negated beneficial effect on the planning. So what does the weight of the

Fig. 9 Difference in transition of the fitness, action steps, and planning limit between the basic model, the nonshared model, and the shared model when the weight of the planning cost was changed. Top row: \( w_p = 0.0 \); bottom row: \( w_p = 0.12 \).

Fig. 10 Transition of the distribution of the planning limit in the population of a certain trial (shared model, \( w_p = 0.12 \)).
planning cost ($w_p$) mean in the evolution of the human intelligence?

The weight of the planning cost ($w_p$) seems to have been changed in the hominids line in several contexts. There seemed to be a change of hominids diet from predominantly vegetarian to more protein and fat based ~2.5 Million years ago because there was discovery of stone tools and associated bones which is the clear evidence that ancestral hominids definitely made sharp-edged cutting tools used for processing carcasses for meat [13]. Eating meat would supply the nutrients to the brain, and then it might exceed the planning cost. Also, recent studies in the brain science have shown that more intelligent people can respond more promptly, precisely, and economically (they consume little glucose) against new and full of variety of situations less intelligent people [14]. Furthermore, Neubaruer et al. (2000) discovered that cognitive processing speed of human is considerably influenced by genetic influence [15]. The influence of cost of planning would be weak if we had a capacity to engage in cognitive task quickly, correctly, and economically.

Considering above factors and results of our simulation, we can present the following scenario as: (1) Tool making intelligence, (2) Social pressure and planning indicate that certain level of ecologic dominance, social competition, and coalitionary arms races: why humans evolved extraordinary intelligence, Evolution and Human Behavior, 26, 10–46, 2005. (8) Alexander R. D., How did humans evolve? Reflections on the uniquely unique species, Museum of Zoology (Special Publication No. 1), Ann Arbor, MI: The University of Michigan, 1990.

Yet, there is still an element of doubt about how hominids could climb a steep cost gradient in the deep environment where the influence of the planning cost was strongly worked. Alexander (1989) argued that once early hominids obtained a certain level of ecological dominance perhaps partly through technological advances (like a stone tool making), then they were faced with increased competition from their own species – "humans uniquely became their own principal hostile force of nature" [16].

The results on social pressure and planning indicate that the planning ability is more adapted to the collective situation by sharing information than that of no sharing information when there is a certain degree of influence of the planning cost. This result implied that there might be a strong connection between the evolution of symbolic communication and planning ability in terms of collective situation.

References