Department of Social Systems and Management

Discussion Paper Series

No.1190

A Simulation Study of Learning a Structure：Mike's Bike Commuting

by

E. AKIYAMA, R. ISHIKAWA, M. KANEKO and J. J. KLINE

February 2008

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

A Simulation Study of Learning a Structure: Mike's Bike Commuting*

E. Akiyama, R. Ishikawa, M. Kaneko† and J. J. Kline‡

February 13, 2008

Abstract

This paper undertakes a simulation study of a player's learning about the structure of a game situation. In a simple 1-person example called Mike's Bike Commuting, we simulate the process by which Mike experiences and accumulates memories about the game structure. It is the basic requirement that to keep an experience as a long-term memory, Mike needs enough repetitions of that experience. By the choice of our simple and casual example, we can discuss relevant time spans for learning. In particular, we argue that the limit case of Mike's learning as time tends to infinity is of little relevance to the problem of learning. We find also that the concept of "marking" introduced by Kaneko-Kline is important for obtaining sufficient structural knowledge in a reasonable time span. The simulation study shows that Mike's learning can change drastically with the concept. We also consider Mike learning his preferences from his experiences, where we will meet various new conceptual problems.

1. Introduction

1.1. Basic Motivations

This paper undertakes a simulation study of a player's learning about the structure of a game situation. It is motivated by the research in inductive game theory, which was initiated by Kaneko-Matsui [14] and Kaneko-Kline [11], [12]. Those papers concentrated

*The authors thank Robert Marks for helpful comments on an earlier version of the paper.

†The first three authors' address: Institute of Policy and Planning Sciences, University of Tsukuba, Ibaraki 305-8673, Japan: These authors are partially supported by Grant-in-Aids for Scientific Research No.18700220, No.17103002, No.19730137, No.1833034, No.19653020, Ministry of Education, Science and Culture.

‡Department of Economics, Bond University, Gold Coarst QDS, Australia: the author is partially supported byAustralian Council Discovery Grant DP0560034.

on the inductive derivation of a personal view from his accumulated memories, without touching on the precise processes of experiencing and accumulating memories. These processes are of a highly finite and complex nature. To consider such complicated but finite problems, we adopt a simulation method. In particular, we can consider the short time span of the social situations; as far as it targets the learning by a human player, the length of time required should be within a player's life time. A simulation study allows us to consider those problems.

In this introduction, we will explain the original motivation for inductive game theory by comparing it with two main stream approaches in the recent game theory literature. They are the classical *ex ante* decision approach and the evolutionary/learning approach. The contrasts between them will be used to motivate our use of a simulation study.

The focus of the classical *ex ante* decision approach is on the relationship between beliefs/knowledge and decision making (cf, Harsanyi [6] for the incomplete information game and Kaneko [10] for the epistemic logic analysis of decision making in a game). In this approach, the beliefs/knowledge are given *a priori* without explicitly formulating their origin, even though their main part must be of an experiential nature.

Contrary to the classical approach, the evolutionary/learning approach, e.g., [21],[5], [9], directs its research target more to experiential worlds. Nevertheless, this approach does not ask the question of the origin/emergence of beliefs/knowledge, either. Instead, its concern is typically the convergence of the distribution of actions to some equilibrium. Now, two points should be emphasized: One is about the term "evolutionary/learning", and the other is about the focus on the limiting (convergence) behavior of a distribution.

The term "evolutionary/learning" means that some effects from past experiences remain in the distribution of genes/actions. It is not about an individual's conscious learning of the game structure; typically it is not even specified who the learner is and what is learned. When we work on an individual's learning, we should make these questions explicit. This requirement leads us to question the focus on the limiting behavior of a distribution: If the learner is an ordinary person, the limiting behavior is not much relevant to his learning. Finiteness of life and learning must be crucial.

Once we take such finiteness and the time span of life seriously, we meet numerous analytical difficulties. Consequently, we conduct simulations over finite spans of time corresponding to the life span, or learning span, of a human being. Our simulation study teaches us that limiting cases are not very relevant to finite learning, as well as more about the types of things a person might be expected to learn in a given finite span of time.

Inductive game theory has aimed to provide answers to the above mentioned questions and criticisms to the extant approaches in the game theory literature. In this paper, we focus on the part of inductive game theory from a player's experiencing to his accumulations of memories. This part is more directly related to the above mentioned questions and criticisms. The relationship between this paper and [14], [11], [12]

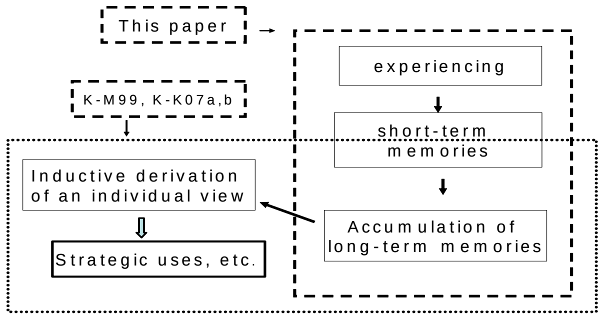

Figure 1.1: Research Map of Inductive Game Theory

is depicted in Fig.1.1. Although these have an overlap, our simulation study is much simpler, from the theoretical point of view, than inductive game theory in [14], [11], [12]. Nevertheless, the study of this paper highlights what kinds of difficulties are involved in accumulation and how we should proceed with our research in inductive game theory.

If one takes the view point, as we do, that game theory should be applied to everyday situations as well, then simulation studies are a method of dealing with the complicated intricacies of those situations. The simulation study helps to point out problems with the extant theories and suggest avenues for further research. From the viewpoint of inductive game theory, the simulation study is still a pilot study and we anticipate significant growth involving further simulation studies. One possibility is to use a computer simulation to study the dynamic changes in an inductively derived view of an individual over time and also interactions between players and their views. In short, we foresee a tremendous variety of problems involved with the process of learning a game structure. We will return to these problems in the conclusion section.

1.2. Simulation Study of the Process of Inductive Learning

We now discuss several points pertinent to inductive game theory and to our basic simulation model. These points help to focus our attention on what we are trying to learn from inductive game theory and our simulation study.

(1): An ordinary person and an every-day situation in a social world: We target the learning of an ordinary human person in a repeated every-day situation, which we regard as being only a small part of the entire social world for that person. To emphasize the ordinariness of the situation, we have chosen a simple and casual example called *"Mike's*

Bike Commuting". This example makes explicit the questions of who the learner is and what is learnt. The learner is Mike, and he learns the various routes to his work. By using an explicit example, the time span and in particular, the number of reasonable repetitions for the experiment also become explicit. The presumption that Mike's bike commuting plays a small part in his life will be used to explain his typical behavior in (3) below.

(2): Ignorance of the structure of the situation: We assume that Mike has no *a priori* beliefs/knowledge about commuting except a minimal map to guide him. At the beginning, Mike's colleague suggested one specific route from his apartment to the office and gave a simple map of possible alternative routes. Everything else about these routes can be learned by Mike only if he takes those routes. It makes sense to ask the question of how many routes Mike is expected to learn after specific lengths of time. Given the finiteness of his life span, we find that the limiting behavior (convergence) of his learning has almost no relevance.

(3): Regular action and occasional deviations: Mike usually follows the suggested route, which we call the *regular route*. Occasionally, though far less frequently, when the mood hits him, he takes a different route.

The assumption of a regular route and occasional deviations is based on (1) combined with the assumption that his energy/time to explore other routes is bounded and scarce. Commuting is only a small part of his entire social world, and he cannot spend his energy/time exclusively exploring those routes. Avoiding deviations may be because of a mental cost such as a fear to see a bad result. This is our starting assumption in this paper, and we do not inquire further into the problem itself.

(4): Short-term memory and long-term memory: We distinguish two types of memories for Mike: short-term and long-term. A short-term memory is a finite time series consisting of past experiences[1], and it will be kept for some finite length of time, perhaps days or weeks. However, when an experience occurs with a certain frequency within his short-term memory, it becomes a long-term memory. Long-term memories are lasting without the time series structure, i.e., neither a sequence of long-term memories nor a precise frequency is recorded.

More generally, short-term and long-term memories come in various forms and with

[1] In the psychology literature, the term "short-term memory" typically means a much shorter memory than what we suppose by the term: for example, the telephone number consisting of 7 digits is kept for one minute in short-term memory. See Deutsch-Deutsch [3]. We have chosen the term "short-term memory" in a way deviant from psychology. This together with Ebbinghous' [4] retention function will be discussed in Section 7. We find many notions of "memories" in the psychology literature such as *episode memory*, *semantic memory* etc. (See Tulving [19], Chap.4 for debates on the various notions of memory). However, with respect to the duration of such a memory, there are only short-term and *long-term* memories. A long-term memory expresses one lasting forever, which is the same as our "long-term memory". Perhaps, the concept of "*everyday memory*" by Linton [8] is the most closely related.

various structures. We emphasize the distinction between short-term and long-term, and formulate them in a specific way. An example given by Linton [7] expresses some structural differences that may exist between short-term and long-term memories, though this example is neither repetitive nor casual. Even now in 2007, most people know about the death of Charles De Gaulle and the landing of the first man on the moon, but very few can recall which came first. However, in 1971, many people could recall which came first. Experiences remain for a certain length of time, but after it, most vanish from our minds. A few remain, but the time series structure connecting them with other experiences may be broken.

In our theory, the transition from a short-term to a long-term memory requires some repetition of the same experience within a short-term memory. This is based on the general idea that memory is reinforced by repetition. Our formulation can be regarded as a simplified version of Ebbinghous' [4] retention function. This will be discussed in Section 7.

(5): Finiteness and complexity: Our learning process is formulated as a stochastic process. Unlike other learning studies, we are not interested in the convergence or any limiting argument. As stated already in (1), the time structure and span are presumed to be finite and short. In the example of Mike's Bike Commuting, we can discuss how many times he has commuted after a half year, one year or ten years. We will find many details, which are highly complex even in this simple example. A simulation study enables us to analyze those details and find the lasting features in Mike's mind.

Even in our simple example, those lengths of time matter significantly and differ from the result for the limit case. Incidentally, we have a convergence result, but use it only as reference.

(6): Marking some actions as important: Although our example is very simple, Mike will find it difficult to fully learn the entire structure of situation even after several years. We consider the positive effect on learning of "marking", introduced in Kaneko-Kline [11]. If Mike marks some actions as "important", and restricts his deviations to the marked actions, then we find that his learning is drastically improved. Without marking, experiences are too infrequent and evaporate with time. By marking, he focuses his attention on fewer experiences, and successfully retains more as long-term memories.

Kaneko-Kline [11] discussed the relation between marking and the size of a derived view. By marking, it would be possible to keep a derived view in a manageable size for the player. Our simulation result indicates that without marking, the view obtained would be extremely poor and small. Thus, our simulation study suggest the importance of marking might not be to obtain a smaller view, but rather a richer view. Imperfections in a player's memory make marking important for obtaining a larger, more informative view.

(7): Learning preferences: So far, our question is about the number of experiences that become long-term memories in a given time span. We extend Mike's learning to his own preferences over routes. Regarding the learning of one's own preferences, we face some new conceptual problems. We mention two important ones.

First, we should make a distinction between having preferences and knowing them. We treat Mike as having well-defined complete preferences, but his knowledge is constrained to only some part by his experiences. Second, learning one's preferences differs from learning information (e.g., from learning the result of an action). Since the feeling of satisfaction is relative and likely to be more transient than the perception of a piece of information, we hypothesize that learning one's preferences needs comparisons close in time. This requirement reduces the number of preferences Mike might expect to learn within a given length of time. Consequently, marking alternatives becomes even more important for obtaining a rich enough view of one's preferences.

This paper is organized as follows: In Section 2, we specify our model and simulation frame. In Section 3, we give simulation results and discuss them to see how much Mike can learn for given time spans. We also explain the structure of our simulation study and the meaning of "probability" within it. In Section 4, we introduce the concept of "marking", and observe it's positive effects on Mike's learning. In Section 6, we carry out a sensitivity analysis of changing various parameters describing Mike's learning and memory characteristics. In Section 5, we consider the problem of Mike learning his preferences. Section 7 is devoted to a discussion our results and their implications for inductive game theory as well as suggesting some future directions for simulations studies.

2. Mike's Bike Commuting

Mike moves to a new town and starts commuting to his office everyday by a bike. At the beginning, his colleague gives him a simple map depicted as Fig.2.1 and mentions the direct route shown by the dotted line. Mike has started commuting every morning and evening, five days a week. From the beginning, he has the intention to know the details of those routes, but the map is abstract and simply symbolic. He decides to explore some routes when the mood hits him, but typically he is busy or tired and simply follows the *regular route* suggested by his colleague[2].

The town has a lattice structure: His apartment and office are located, respectively, at the south-west and north-east corners. To have a direct route from his apartment to the office, he should choose "North" or "East" at each lattice point. There are 35 direct

[2]We may start with only the assumption that he is given the regular route, without having a map. This case is more faithful to inductive game theory in Kaneko-Kline [11], [12]. However, this makes our simulation study much more complicated. We will keep our study as simple as possible.

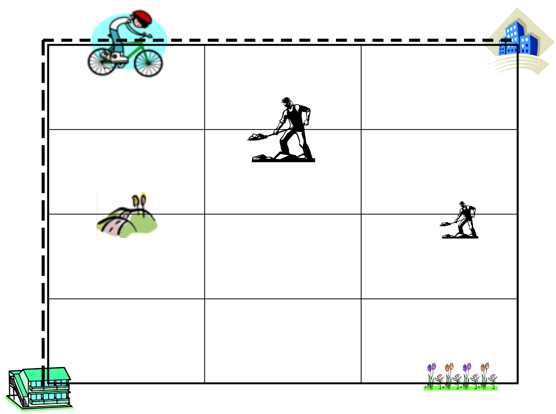

Figure 2.1: Map of Routes

routes. He enumerates these routes as $a_0, a_1, ..., a_{34}$, where a_0 denotes the regular route.

In our simulation, we assume that Mike follows his regular route a_0 with probability $4/5 = 1-p$ and he makes a deviation to some other route with $p = 1/5$. This probability p is called the *deviation probability*. When he makes a deviation, he chooses one route from the remaining 34 routes with the same probability $1/34$. His behavior each morning or evening can be depicted by the tree in Fig.2.2. He himself may not be conscious of these probabilities or of this tree.

Now, we turn our attention to the description of Mike's memory ability. As mentioned in Section 1.2, he has two types of memories: *short-term* and *long-term*. A short-term memory is a time series of experiences of the past m trips. An experience disappears after m trips of commuting. If the same experience, say a_l, occurs at least k times, experience a_l becomes a long-term memory. Long-term memories form a set of experiences without time-structure. In Section 1.2, we explained why long-term memories might not have a time structure or a frequency.

In our main simulation study, we specify the parameters (m, k) as $(10, 2)$, meaning that Mike's short-term memory has length 10, and if a specific experience occurs at least two times in his short-term memory, it becomes a long-term memory. This situation is depicted in Fig.2.3, where at time $t-1$, the routes a_0, a_2 are already long-term memories, and at time t, route a_1 becomes a new long-term memory.

In our simulation study, we consider one more parameter T, denoting the total number of trips (time span). For example:

$$\text{after a half year, } T \;=\; 2 \times 5 \text{ (days)} \times 25 \text{ (weeks)} = 250;$$

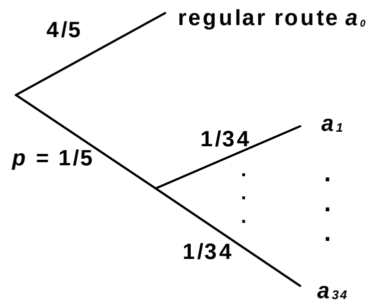

Figure 2.2: Mike's behavior in each trip of commuting

$$\begin{aligned} \text{after 1 year, } T &= 2 \times 5 \text{ (days)} \times 50 \text{ (weeks)} = 500; \\ \text{after 10 years, } T &= 2 \times 5 \text{ (days)} \times 500 \text{ (weeks)} = 5000. \end{aligned}$$

Our simulation will be done by focussing on the half year and 10 year time spans. From time to time, however, we will refer to some other time spans.

In Mike's Bike Commuting, the number of available routes is 35. Later, this parameter will also be changed, and the number of routes will be denoted as a parameter s. Listing all the parameters, we have our *simulation frame* F:

$$F = [s, p; (m, k)]. \tag{2.1}$$

We will always assume that in the case of a deviation, a route other the regular one is chosen with equal probability $1/(s-1)$.

The stochastic process is determined by the simulation frame F when T is given, which consists of T component stochastic trees depicted in Fig.2.2. This stochastic process is denoted by $F[T] = [s, p; (m, k) : T]$. Our concern is the probability of some event related to long-term memories at time T. For example, at T, what is the probability of the event that a particular route a_l is a long-term memory? Or what is the probability that all routes are long-term memories? We calculate those probabilities by simulation. In Section 3, we will give our simulation results for $F = [s, p; (m, k)] = [35, 1/5; (10, 2)]$ and $T = 250$, 5000. We will also explain the meaning of "probability" in our simulation study.

Before going to these results, we mention one analytic result: For the stochastic process $F[T] = [35, 1/5; (10, 2) : T]$,

$$\text{the probability that all routes become long-term memories} \tag{2.2}$$

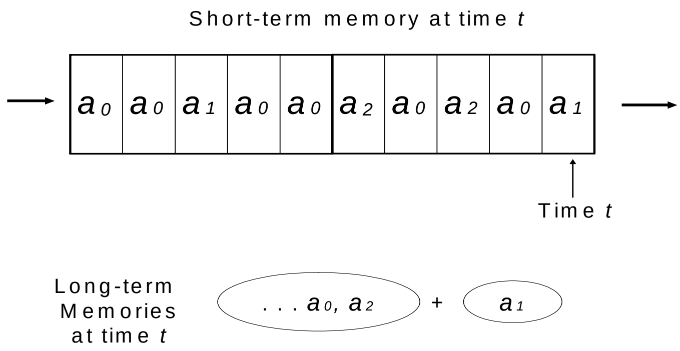

Figure 2.3: Short-term memory and long-term memory

tends to 1 as T tends to infinity.

This can be proved easily because the same experience occurs twice in a short-term memory at some point of time almost surely if T is unbounded. This result does not depend on the specification of parameters of F. We should ask the question: how relevant is this convergence result to finite cases such as $T = 250$ and $T = 5000$? Our finding by simulation is that the convergence result gives no specific information about the cases of $T = 250$ and $T = 5000$.

3. Simulations for $s = 35$ and the Meaning of "Probability"

In this section, we give simulation results in the case of $s = 35$. The results confirm that it would be difficult for Mike to learn all the routes after a half year. After ten years, he learns more routes, but we cannot say much about which specific routes he learns other than the regular one. In Section 3.2., we will explain our simulation method and the meaning of "probability".

3.1. Simulation Results for $s = 35$

Consider the stochastic process determined by $F = [s, p : (m, k)] = [35, 1/5; (10, 2)]$ up to $T = 250$ and 5000. Table 3.1 provides the probabilities of the event that a specific route a_l is a long-term memory at $T = 250$, and also at quite a large T which will be explained shortly:

Table 3.1

T	250	5000	28252 ($>$ 56 years)
a_0	1	1	1
a_l ($l \neq 0$)	0.069	0.765	0.99

The row for a_0 shows that the probability of the regular route a_0 being a long-term memory is already 1 even at $T = 250$ (a half year). This "1" is, in fact, an approximation result, since analytically the event that the regular route is not a long-term memory has a very small but positive probability. We have put a "1" there simply because the regular route is a long-term memory in all of our simulation runs even at $T = 250$. From now on, we regard the probability of a_0 being a long-term memory simply as 1 for $T \geq 250$.

The row for a_l ($l \neq 0$) is more interesting. The probability that a specific a_l ($l \neq 0$) is a long-term memory at $T = 250$ and 5000 is 0.069 and 0.765, respectively. Our main concern is to evaluate these probabilities from the viewpoint of Mike's learning.

Some reader may have expected that for $T = 250$ the probability would be much smaller than 0.069, because in each trip, the probability of route a_l ($l \neq 0$) being chosen is only $1/5 \times 1/34 = 1/170 = 0.00588$. However, it is enough for a_l to occur in a consecutive sequence of length 10 (short-term memory) at some $t \leq 250$, and there are 240 such consecutive ten sequences. Hence, the probability turns out not to be negligible[3]. The accuracy of this calculation will be discussed in Section 3.2.

The rightmost column is prepared for a purpose of reference. The number of trips 28252 ($>$ 56 years) is obtained from asking the time span needed to obtain the probability 0.99 of a_l ($l \neq 0$) being a long-term memory. The length of 56 years would typically exceed an individual career,[4] and thus we regard the limiting result (2.2) as only a reference.

The cases of $T = 250$ and 5000 are more relevant to our analysis. Nevertheless, a single probability 0.069 or 0.765 tells us little about what Mike might be expected to learn in those time spans. We next look more closely at the distribution of routes he learns for each of those time spans.

For $T = 250$ we give Table 3.2, which describes the probability of exactly r routes (the regular route and $r - 1$ alternative routes) being long-term memories:

[3] A famous example called the *birthday attack* may be indicative for this fact: In a class consisting 50 students, what is the probability of finding at least one pair of students having the same birthday? Since each student has the probability 1/365 of an arbitrary given day of a year being his birthday, it might be expected to have a pair of students of the same birthday. However, the exact calculation tells that the probability is about 0.97.

[4] Our model without decay of long-term memories is likely to be inappropriate for 56 years.

Table 3.2

r	1	2	3	4	5	$\cdots$
	0.089	0.223	0.272	0.213	0.121	$\cdots$

After $r = 5$ routes the probability is diminishing quickly so we exclude those numbers from the table. According to our results, Mike typically learns a few routes (the average is about 3.33) after half a year. However, since there are 35 routes in total, there are many cases to be considered. For $r = 3$, one route must be regular but others two are arbitrary. We have $\binom{34}{2} = 561$ cases, so the probability of a particular 3 routes being long-term memories is only $0.272/561 = 0.000485$ which is quite small. This means that although Mike learns about 2 alternative routes, it is hard to predict with much accuracy which pair would be learned.

As time goes on, Mike continues his exploration of those routes. At $T = 5000$, i.e., ten year later, Mike's learning is described by Table 3.3.

Table 3.3

r	$\cdots$	25	26	27	28	29	$\cdots$	
	$\cdots$	0.109	0.159	0.153	0.153	0.124	$\cdots$	

Again, we show only the values of r having high probabilities. The average of the number of routes as long-term memories is about 27.0. Because most of the distribution lies between 25 and 29 routes, we find that there many more cases to consider than after half a year. For example, consider 0.109 for $r = 25$, which is the probability that exactly 25 routes are learned. Incidentally, this probability can be obtained from the probability 0.764 in Table 3.1 by the equation:

$$\binom{34}{24} \times (0.764)^{24} \times (1 - 0.764)^{9} \doteqdot 0.109.$$

Looking at this equation, we obtain the probability that a specific set of 25 routes are long-term memories is only $0.109/\binom{34}{24} = 8.31 \times 10^{-10}$. In sum, Mike learns about 27.0 alternative routes after 10 years. However, the number of combinations of 24 routes from 34 is enormous at about 1.3×10^{8} and much larger than the $\binom{34}{2} = 561$ cases we need to consider after only half a year. While Mike learns many more routes after 25 years, it is almost impossible to predict which combination of 25 routes he will learn. If we focus on the limiting cases, then Mike learns all routes almost surely. However, for more realistic time spans, we cannot predict well what he will know even after quite long spans.

Finally, we report the average time for Mike to learn all the 35 routes as long-term memories, which is 28.4 years $(14,224.3$ trips). If he is very lucky, he will learn all routes in a short length of time, say, 10 years, which is an unlikely event of probability 9×10^{-5}. The probability of having learned all routes in 35 years is much higher at 0.806.

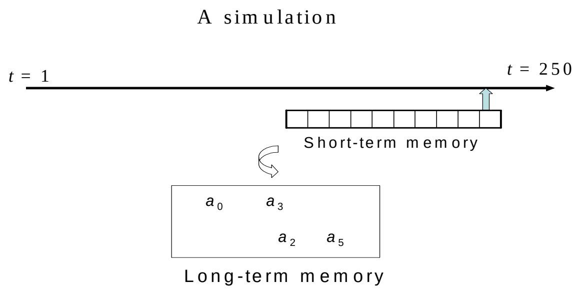

Figure 3.1: a simulation up to 250

3.2. Simulations and the Meaning of "Probability"

We now explain the concept of "probability" we are using, and discuss the accuracy of this concept. First we mention why this is not calculated in an analytic manner. The analytic computation is feasible up to about $T = 30$, and beyond $T = 40$, it is practically impossible in the sense that for $T = 50$, it takes more than 170 years by the current computer ability with our analytical method. The reason for this complication in analytical computation is caused by a limited length of short-term memory and multiple occurrences needed for a long-term memory.

Instead of using an analytic method, we take the relative frequency of a given event over many simulation runs. We use the Monte Carlo method to simulate the stochastic process up to a specific T for the simulation frame $F = [s, p : (m, k)] = [35, 1/5 : (10, 2)]$. The frame has only two random mechanisms depicted in Fig.2.2, but they are reduced into one random mechanism. This mechanism is simulated by a random number generator. Then, we simulate the stochastic process determined by F up to $T = 250$ or $T = 5000$ or some other time span. A simulation is depicted in Fig.3.1. One simulation run gives a set of long-term memories: In Fig.3.1, routes a_0, a_2, a_3, a_5 are long-term memories at some time before $T = 250$.

We run this simulation $100,000$ times. The "probability" of a_l is calculated as the relative frequency:

$$\frac{\#\{\text{simulation runs with } a_l \text{ as a long-term memory}\}}{100,000} \tag{3.1}$$

In the case of $T = 250$, this frequency is about 0.069 for $l \neq 0$, and it is already 1 for $l = 0$ in our simulation study (see Section 3.1).

We compare some results from simulation to the results by the analytical method. For $T = 20$ and $s = 35$, the probability of a_l being a long-term memory can be calculated in an analytic manner using a computer. The result coincides with the frequency obtained using simulation to accuracy of 10^{-4}.

The accuracy of the frequency (probability) 0.069 in Table 3.1 is evaluated further by looking at $1,000,000,000$ simulation runs. In these runs, we have $68,594,265$ runs where a_1 is a long-term memory. Counting also simulation runs where a_l $(= a_2, ..., a_{34})$ is a long-term memory, we find that the smallest (and largest) number of runs where a_l is a long-term memory is $68,569,941$ (respectively, $68,596,187$). We have the interval from 0.06857 to 0.06860, both of which become 0.069 by rounding them off to three decimal places.

In sum, we mainly calculate the "probability" of an event as the relative frequency over numerous simulation runs since the analytic calculation is not feasible for the time spans and simulation frames under consideration. We test the accuracy of our definition by both comparing it to analytic results for shorter time spans, and by showing that the numbers are robust with respect to changes in the number of simulation runs.

4. Learning with Marking: Simulation for $s = 5$

In this section, we discuss how "marking", introduced in Kaneko-Kline [11], can improve Mike's learning. By concentrating his efforts on a few "marked" routes, he is able to learn and retain more experiences. This is because the likelihood of having an experience repeated rises when he reduces the number of experimental alternatives. Without marking, each experience occurs too infrequently and evaporates from his mind before it becomes a long-term memory. In Section 4.1, we consider the case where Mike marks only four alternative routes in addition to the regular one. We will see a dramatic effect in his learning of alternative routes coming from this simple change in his behavior. In Section 4.2, we show how a more planned approach can improve the effect of "marking" on his learning.

4.1. Marking Five Routes and Simulation Results

Suppose that Mike decides to mark some routes from the map as more important for his exploration. He uses two criteria to choose a deviation route: (i) he chooses pleasant routes like those having a scenic hill or flowers; and (ii) he avoids construction sites. Then, he finds only four alternative routes worth trying, which are depicted in Fig.4.1. Adding the regular route a_0 as marked too, we denote the five marked routes by a_0, a_1, a_2, a_3, a_4.

The above situation is described by changing the previous simulation frame to $F = [s, p : (m, k)] = [5, 1/5 : (10, 2)]$ for $T = 250$ or 5000, where $s = 35$ changes into $s = 5$

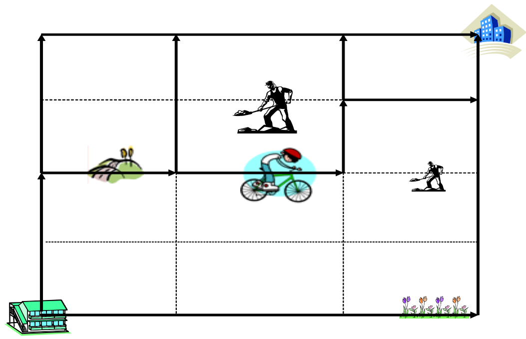

Figure 4.1: Five Marked Routes

only. We can calculate, by our simulation method, the probability of a_l ($l \neq 0$) being a long-term memory, which is given in Table 4.1:

Table 4.1

T	250	5000
$s = 5$	0.970	1.00
$s = 35$	0.069	0.765

Table 4.2

T	425	28253
$s = 5$	0.990	1.000
$s = 35$	0.114	0.990

The probability of learning a route even after a half year is now drastically improved by this marking. Table 4.2 gives the length of time required for the probability of a_l ($l \neq 0$) being a long-term memory to be 0.99. With marking $s = 5$, he needs only 425 trips (10.2 months), as opposed to the 28, 253 trips (more than 56 years) in the case of $s = 35$.

We also have calculated the probability that exactly r number of routes are long-term memories at $T = 250$, which is given in Table 4.3.

Table 4.3

r	1	2	3	4	5
	8.00×10^{-7}	1.04×10^{-4}	5.05×10^{-3}	0.109	0.886

In this case, the average number of routes learned is 4.9. Incidentally, as reported in Table 4.4, the average time for Mike to learn all 35 routes without marking is about 100 times the average time to learn 5 routes by marking. This suggests that Mike might be able to use marking in a more sophisticated manner to learn all 35 routes in a shorter period of time than the 28.4 years required without marking. We will look more closely at this idea in Section 4.2.

Table 4.4

	$s = 5$	$s = 35$
the average number of trips to learn all	151.8 3.6 months	$14,224.3$ 28.4 years

4.2. Learning by Marking and Filtering

Suppose that Mike has learned all four marked alternative routes in addition to the regular route after a half year. He may then want to explore some other routes. He might plan to explore the other 30 routes by dividing them into 6 bundles of 5 routes, trying to learn each bundle one by one. We suppose that he explores one bundle for a half year, and he moves to the next bundle storing any long-term memories in the process. Thus, Mike has discovered a method of filtering to improve his learning.

According to the result of Section 4.1, Mike most likely learns all five routes within a half year. By his filtering he reduces the expected time to learn all 35 routes from 28.4 years to only $250 \times 7 = 1750$ (3.5 years).

We can simulate the above process and calculate the probability of that he finishes his entire exploration in 3.5 years. This probability is $(0.886)^7 \doteqdot 0.427$, and with the remaining probability 0.573, at least one route is not learned after 3.5 years. If some routes still remain unlearned, then we assume that he rebundles the remaining routes into bundles of 5. However, we expect rather small number of unlearned routes to remain; the event of 3 remaining is rare event occurring with only probability 0.03. Most likely, Mike's learning finishes within 4 years.

The results of this section have the implication that the use of "marking" together with filtering can be very useful for learning, while without marking, it would become very difficult as the number of routes becomes slightly bigger.

If we treat the above filtering method alone, forgetting the original constraint such as the energy-scarcity mentioned in Section 1.2, the extreme case would be that he chooses and fixes one route for two trips and goes to another route. In this way, he could learn all routes with certainty in precisely 35 days. However, this type of short-sighted optimal programming goes against our original intention of exploration being rather rare and unplanned. Commuting is one of many everyday activities for Mike, and he cannot spend his energy/time exclusively on planning and undertaking this activities. Though our example is very simplified, we should not forget that many unwritten constraints lie behind it, which are still significant to Mike's learning.

5. Learning Preferences

In this section, we consider Mike's learning of his own preferences. We treat Mike's preferences and satisfaction as being based on his comparisons between experienced routes. We do not start with a numerical representation of his preferences such as a utility function. Mike needs to find his own preferences by his experiences. Now, we meet various conceptual problems. First, we should specify the bases for our analysis, and then we formulate the process by which Mike learns his own preferences. We simulate this learning process in Section 5.1, and show that learning of his preferences is typically much slower than learning routes. Consequently, notions like "marking" become even more indispensable. In Section 5.2, we consider the change of the process when he adopts a more satisfying route based on his past experiences.

5.1. Preferences

In our simulation of Mike's learning, each route was treated as just a symbol without any additional structure. There are various additional structures we can look at, for example, the internal structure (substructure) of details he finds along each route. Alternatively, we can add a superstructure such as preferences over routes. We pursue this superstructure in the present section.

Since Mike has no idea of details along each route at the beginning, one might wonder if he has well-defined preferences over the routes or what form they would take. However, by looking at preferences in their original meaning, we can connect them with experiences. Since an experience of each route gives some level of satisfaction, comparisons between such satisfaction levels can be regarded as his preferences. Here, preferences may be regarded as being inherent, but they are only revealed to Mike himself when he experiences and compares different outcomes. In this way, Mike may come to know some of his own preferences.

A preference between two routes is experienced only by comparing the two satisfaction levels from those routes, which should be distinguished from the notion of "revealed preferences" (cf. Malinvoud [18])[5]. A feeling of a satisfaction typically emerges in the mind (brain) without tangible pieces of information. Such feelings may often be transient and only remain after being expressed by some language such as "this wine is better than yesterday's". We assume, firstly, that "satisfaction" is of a transient nature, and secondly, that the satisfaction from one event can be compared with that of another event only if these events are close enough in time.

In this paper, we assume that Mike has a complete and transitive preference relation

[5] It is the view that a preference is defined by a (revealed) choice from hypothetically given two alternatives. It is our point that this hypothetical choice is highly problematic from the experiential point of view.

over the routes, which he himself does not know. We do not ask how his preferences are formed and/or measured[6], but we ask how much he will learn about his own preferences by experiencing them.

In our simulation study, we formulate a preference comparison between a pair of routes experienced in consecutive trips as an experience. However, this experience has a quite different nature from a sole experience of a route. The former needs the comparison of two experienced satisfaction levels. To distinguish between these different types of experiences, we call a sole experience of a route a *first-order experience*, while a pairwise comparison of two routes is a *second-order experience*. Second-order experiences are of a different nature and should be treated differently[7].

Consider Mike's learning of such second-order experiences in the simulation frame $F = [s, p : (m, s)] = [5, 1/5 : (10, 2)]$ and $T = 250$ or 5000. A short-term memory is now treated as a sequence of length 10, where each element in the sequence is an unordered pair of consecutively experienced routes. Consecutive routes can be compared to form preferences over the pair. For example, in Table5.1, the short-term memory is the sequence of 10 unordered pairs $\langle a_1, a_0\rangle, \langle a_0, a_0\rangle, ..., \langle a_{3,} a_0\rangle$. From the pair $\langle a_1, a_0\rangle$, for example, Mike finds the second-order experience of a preference comparison between a_1 and a_0. These second-order experiences are what may become long-term memories.

Table 5.1

a_1	$^{a_1}\ a_0$	$^{a_0}\ a_0$	$^{a_0}\ a_0$	$^{a_0}\ a_0$	$^{a_0}\ a_1$	$^{a_1}\ a_2$	$^{a_2}\ a_0$	$^{a_0}\ a_0$	$^{a_0}\ a_3$	$^{a_3}\ a_0$
$\rightarrow$	$t-9$	$t-8$	$t-7$	$t-6$	$t-5$	$t-4$	$t-3$	$t-2$	$t-1$	$t \ \rightarrow$

For a second-order experience to become a long-term memory, however, it must occur at least twice in a short-term memory. In Fig.5.1, $\langle a_0, a_1\rangle$ occurs twice at trips $t-9$ and $t-5$, and hence it becomes a long-term memory. We assume that these unordered pairs are disjoint, and remark that $\langle a_3, a_0\rangle$ is not counted twice in the short-term memory at trips $t-1$ and t, since a_3 itself is experienced only once at $t-1$.

The computation result is given for $l, l' = 1, 2, 3, 4$ and $l \neq l'$ in Table 5.2.

[6] For philosophical discussions on these problems, see von Neumann-Morgenstern [22] and Luce-Raiffa [17].

[7] We borrow these terms from mathematical logic, since the distinction is conceptually similar. Nevertheless, our use of these terms slighlty differs from in mathematical logic. In mathematical logic, a variable representing a pair of objects can be treated in a first-order theory. The second-order concepts are really necessary for a variable to represent some structure including a infinite number of objects. See Kleene [15].

Table 5.2

trips	Prob. of comparison a_0 vs. a_l	Prob. of comparison a_l vs. $a_{l'}$
250 (a half year)	0.981	0.053
5000 (10 years)	1.000	0.671
10000 (20 years)	1.000	0.892

In the column of a_0 vs. a_l, the probability of the preference between a_0 and a_l being a long-term memory is given as 0.981 for $T = 250$. After 10 years and 20 years, the probabilities are already 1; actually, it becomes already 1 after only about 2 years.

One might expect that the value of 0.981 for a comparison between a_0 and a_l ($l \neq 0$) would be lower than the 0.970 for just learning a route a_l reported in Table 4.1. This discrepancy can be explained by the counting of pairs at the boundary. For example, the comparison between a_0 and a_1 appearing in Table 5.1 becomes a long-term memory from the short-term memory at time t. However, in our previous treatment of memory of routes, a_1 would not be a long-term memory.

When it comes to comparing two alternative routes, we find in the right column of Table 5.2 that the learning is very slow. After a half year, Mike hardly learns any of his preferences between alternative routes. An experience of comparison between a_l vs. $a_{l'}$ happens with such a small probability, because both deviations a_l and $a_{l'}$ from the regular route a_0 are required consecutively and also twice separately. This means that his learned preferences are very *incomplete* after a half year.

Let us look more closely at what is happening for $T = 250$. Since the probability of two alternative routes being a long-term memory is very small, it would be more informative to focus only on a_0 vs. a_l for $l = 1, 2, 3, 4$. Table 5.3 gives the probability of the event that exactly r preferences between a_0 and a_l ($l \neq 0$) are long-term memories:

Table 5.3

r	0	1	2	3	4
prob.	1.30×10^{-7}	2.69×10^{-5}	2.08×10^{-3}	7.17×10^{-2}	0.926

This shows that almost certainly the preference between the regular behavior and each of the 4 alternatives is known within half a year. On the other hand, the preference between two alternatives is rarely a long-term memory after half a year.

It is an important implication that Mike's learned preference relation is far from being *complete*, though his original (hidden) preference relation is complete. For example, suppose that Mike's original preference relation is the strict order, a_3, a_4, a_0, a_1, a_2 with a_3 at the top. After half a year, he likely learns his preferences between a_0 (regular) and each alternative $a_l, l = 1, 2, 3, 4$, which is depicted in the middle diagram of Table 5.4. It is unlikely that he learns which of a_3 or a_4 (or, a_1 or a_2) is better. Even if he believes *transitivity* in his preferences, he would only infer from his learned preferences that both a_3 and a_4 are better than a_1 and a_2. Without transitivity, even this fact could not be derived.

Table 5.4

$$
\begin{array}{l}
a_3 \\
a_4 \\
\mathbf{a}_0 \text{ (reg.)} \\
a_1 \\
a_2
\end{array}
\Longrightarrow
\begin{array}{ccc}
a_3 & & \mathbf{a}_4(\text{reg.}) \\
\nwarrow & & \nearrow \\
& a_0 & \\
\nearrow & & \nwarrow \\
a_1 & & a_2
\end{array}
\Longrightarrow
\begin{array}{ccc}
& \mathbf{a}_3 \text{ (reg)} & \\
& \uparrow & \\
& a_4 & \\
& \uparrow & \\
& a_0 & \\
\nearrow & & \nwarrow \\
a_1 & & a_2
\end{array}
$$

Ten years later, Mike's knowledge will be much improved. By this time, with probability 1, he will have learned his preferences between a_0 and each alternative $a_l, l = 1, 2, 3, 4$. He will also likely have learned his preferences between some of the alternatives. Table 5.5 lists the probabilities that exactly r of his preferences are learned. Recall that there are $\binom{5}{2} = 10$ comparisons. We see that even after 10 years, Mike is still learning his own preferences over alternative routes.

Table 5.5 (10 years)

r	4	5	6	7	8	9	10
	1.07×10^{-3}	0.0155	0.079	0.215	0.329	0.269	0.0913

However, after 20 years, he learns much more about his preferences, which is described in Table 5.6.

Table 5.6 (20 years)

r	4	5	6	7	8	9	10
	1.59×10^{-15}	7.86×10^{-5}	0.0016	0.0179	0.111	0.366	0.504

As it happens, by the time Mike is able to get to taste the rough with the smooth, he is already old.

5.2. Maximizing Preferences

The results of the previous subsection tell us that it is very difficult for Mike to learn his complete preferences. However, completeness should not be his concern. For him, it would be important to find a better route than the regular one, and to change his regular behavior to the best route he knows. This idea is formulated as follows:

(1): he continues to learn his preferences until he can compare each marked alternative to the regular one;

(2): if finds a better route a_l than a_0 in those comparisons, then he would choose a_l as the new regular route;

(3): he stores a_0 and the alternative routes less preferred than a_0;

(4): he makes an exploration of his preferences over the remaining marked alternatives with the new regular route a_l;

(5): he repeats the process determined similarly by $(1)-(4)$ until he does not find a better route than the regular one[8].

When there are multiple better routes in (2), the choice of a new regular route is assumed to be arbitrary. Apparently, the final result of the process determined by $(1)-(5)$ gives a highest preference. Our concern is the length of time for this process to finish, and his knowledge about his preferences upon finishing.

Suppose that Mike's original (hidden) preferences are described by the left column of Table 5.4; he has a strict preference ordering $a_3 \prec a_4 \prec a_0 \prec a_1 \prec a_2$, where a_0 is the regular route. Initially, he knows nothing about his preferences. After some time, he learns his preferences described in the middle diagram. In this case, very likely, only his preferences between a_0 vs. a_l $(l \neq 0)$ are learned. The thin arrow $\rightarrow$ indicates the learned preferences.

We could assume as in Section 4.2 that Mike changes his behavior after a half year. But here, let us see the average time to finish his learning for preference maximization, under the *assumption* that as soon as he finishes his learning of the preferences between the regular route and alternative ones, he moves to learning the unlearned better part. The transition from the left column to the middle one in Table 5.4 needs the average time 136.2 (3.3 months).

When he reached the middle diagram of Table 5.4, he starts comparing between a_3 and a_4. Here, a_4 is taken as the new regular route. After the middle diagram, he stores the preferences over a_0, a_1 and a_2. Once he obtains the preference between a_3 and a_4, he goes to the right diagram and he plays the most preferred route a_3. The average time for this second transition is 11.0 trips (1.1 week). Hence, the transition from the left diagram of knowing no preferences, to the right diagram of having a preference maximizing route takes the average time $136.2 + 11.0 = 147.2$ trips (3.5 months).

After the above learning, he has learned only some part of his entire preferences, e.g., he has no idea the preference between the bottom two routes. If he uses transitivity, he can compare the top to every alternative route, but his knowledge of preferences is still incomplete. Transitivity is needed for Mike to ensure that a_3 is the top for him.

In the above calculation, we started with the preference ordering given in the left diagram of Table 5.4. For all possible preference orderings, we have $5! = 120$ cases, but can classify them into 5 classes by the position of a_0. Here we consider only the

[8] Exactly speaking, this process should be inductively defined, but we avoid it to keep a simple explanation.

two cases where a_0 is the top or the bottom. When a_0 is the top, only one round is enough to learn this fact, which takes the average time 136.2 (3.3 months). However, Mike learns no other preferences.

When a_0 is the bottom, there are several cases depending upon his choice of new regular routes. But now there are four possibilities for the choice of the next regular route. Depending upon this choice, he may finish quickly or needs more rounds. The more quickly he finishes, the more incomplete are his preferences. Alternatively, the slowest case for finding the top needs 4 rounds. Table 5.7 depicts the slowest case (the total average time is $136.2 + 78.0 + 36.4 + 11.0 = 261.6$ (6.3 months)), where the bold letter means the regular route. By this process he recovers his own complete preferences, still, with the help of transitivity.

In Sum, if Mike learns the top quickly, he learns virtually nothing about his preferences between the other alternatives. If he finds the top slowly, however, he would have a much richer knowledge of his own preferences[9].

Table 5.7: Transitions with learning preference

$$\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \mathbf{a}_0 \end{array} \Longrightarrow \begin{array}{cccc} a_1 & a_2 & a_3 & \mathbf{a}_4 \\ \nwarrow & \uparrow & \uparrow & \nearrow \\ & a_0 & & \end{array} \Longrightarrow \begin{array}{ccc} a_1 & a_2 & \mathbf{a}_3 \\ \nwarrow & \nwarrow & \nearrow \\ & a_4 & \\ & \uparrow & \\ & a_0 & \end{array} \Longrightarrow \begin{array}{cc} a_1 & \mathbf{a}_2 \\ \nwarrow & \uparrow \\ a_3 & \\ \uparrow & \\ a_4 & \\ \uparrow & \\ a_0 & \end{array} \Longrightarrow \begin{array}{c} \mathbf{a}_1 \\ \uparrow \\ a_2 \\ \uparrow \\ a_3 \\ \uparrow \\ a_4 \\ \uparrow \\ a_0 \end{array}$$

6. Sensitivities with Parameter Changes

Mike's learning is described by the stochastic process $F[T]$ determined by a simulation frame $F = [s, p; (m, k)]$. So far, we have seen the effects of changes of s and T on Mike's learning. In this section, we briefly consider the sensitivity of Mike's learning to the other parameters p (deviation probability), m (length of a short-term memory), k (threshold number).

The deviation probability p and the other two (m, k) are of a different nature. While the deviation probability p is regarded as externally given it may be controlled by Mike

[9]Some reader may wonder what implications this argument has on the discounted sum of future utilities. Even under the stationarity assumption that preferences are time independent, this problem of time preferences requires the 3rd-oder experiences, i.e., a preference between a present outcome and a next outcome should be compared with another preference. Without the stationary assumption, experiences of any orders are required. In this sense, from the experiential point view, the discounted sum of future utilities are out of the scope.

in an effort to learn more about alternative routes. We should keep in mind that our intention is to capture casual everyday learning. Raising p would require additional effort and and likely have significant opportunity costs. Thus we should keep the values of p low. The parameters m and k may also be within Mike's control, but because they describe his memory ability, changing them may also require greater effort on his part. Whether or not these are in Mike's control, it is still interesting to find out how sensitive his learning is to these parameters.

We start with a sensitivity analysis of learning to changes in m and k. Let $p = 1/5$ and $s = 5$. Table 6.1 gives the probability of a specific route a_l ($l \neq 0$) being a long-term memory for the cases of $k = 1, 2, 3$ with $m = 10$. Focusing on $T = 250$, the drop in probability from 0.970 for $k = 2$ to 0.488 for $k = 3$ suggests that Mike's learning is quite sensitive to changes in k.

Table 6.1 : $s = 5$ and $m = 10$

	$T = 250$	$T = 5000$
$k = 1$	1.000	1.000
$k = 2$	0.970	1.000
$k = 3$	0.488	1.000

Table 6.2 : $s = 5$ and $k = 2$

	$T = 250$	$T = 5000$
$m = 7$	0.930	1.000
$m = 10$	0.970	1.000
$m = 20$	0.995	1.000

On the other hand, Table 6.2 suggests that his learning is less sensitive to the change in the length m of each short-term memory.

When m and k change simultaneously for $s = 5, 35$, we have the results listed in Tables 6.3 and 6.4.

Table 6.3 : $s = 5$

(m, k)	$T = 250$	$T = 5000$
$(10, 2)$	0.970	1.000
$(20, 3)$	0.840	1.000

Table 6.4 : $s = 35$

(m, k)	$T = 250$	$T = 5000$
$(10, 2)$	0.068	0.765
$(20, 3)$	0.007	0.140

Table 6.4 shows that increasing both k and m implies that Mikes learning can also be affected a lot. In the case of $s = 35$ his learning of a single alternative becomes much worse. However, from Table 6.3, we find the implication that "marking" still helps Mike a lot.

Finally, we consider how sensitive Mike's learning is with respect to the probability of deviations p. We look at how his learning changes when p changes from 1/5 to 0.05, 0.1 and 0.3. We focus on the probability that a specific a_l ($l \neq 0$) becomes a long-term memory for the cases of $s = 5$, 35 and $T = 250$, 5000. The results are given in Tables 6.1 and 6.2:

Table 6.5 : $s = 5$

$p \setminus T$	250	5000	Av.no
0.05	0.259	0.998	1720
0.1	0.655	1.000	488.6
0.2	0.970	1.000	151.7
0.3	0.999	1.000	80.24

Table 6.6 : $s = 35$

$p \setminus T$	250	5000	Av.no
0.05	0.005	0.091	215707
0.1	0.018	0.312	54893
0.2	0.068	0.765	14223
0.3	0.143	0.957	6548.5

We find that the probability of a_l ($l \neq 0$) being a long-term memory is quite sensitive to a change in p. In the case of $s = 5$, when $p = 0.1 = 1/10$ or $0.05 = 1/20$, the probability of an alternative route becoming a long-term memory after a half year is much smaller than at $p = 1/5$. In the case of $s = 35$, the decrease in this probability is even more dramatic. On the other hand, increasing p to 0.3 has quite a large effect of raising the probability to almost 1 even for half a year. The right-most columns of Tables 6.5 and 6.6 also list the average number of trips needed to have all routes being long-term memories. These numbers are seen to also be highly sensitive to changes in p.

The changes of deviation probability p should be interpreted while taking (1) of Section 1.2 into account. That is, if commuting is a small part of his entire social world, then p should be a relatively small value such as 0.2 or 0.05. If Mike is not busy with other work, and he keeps enough energy and curiosity about details of the routes, it may be as high as 0.3. On the other hand, 0.3 means that he uses his energy three times in a week, and his behavior may be interpreted as shirking by his boss.

7. Concluding Discussions

In this paper, we undertook a simulation study of one-person learning about the structure of an everyday situation. The study has shown that how finite learning may differ substantively from both our expectations and also from the limit case of unbounded learning. We summarize our main findings and emphasize their implications in this section. We also mention some avenues for further developments of simulation studies and for inductive game theory.

The example "Mike's Bike Commuting" is a small everyday situation. It therefore provides insights to our everyday lives and everyday behavior. Since the life span of a human being has a definite upper bound, we can explicitly formulate and compute what learning is possible and relevant to a person given his life span.

When we consider the implications of our simulation results, we should not forget the underlying background of our study. Recall that our target situation is always partial relative a player's entire social world. This explains the regular behavior as a method of time/energy saving and at the same time infrequent deviations as an exploration behavior. Once the problem is mathematized explicitly, we might have a tendency to study its theoretical possibilities as if it is a closed world. One example of this kind was mentioned in the last paragraph in Section 4.2.

We can also consider the implications of our study to game theory. Our original motivation was, from the viewpoint of inductive game theory, to study the origin/emergence of beliefs/knowledge of the structure of the game. Long-term memories are the source for such beliefs/knowledge. Our results have the implication that it would be difficult for a player to learn the full structure of a game unless it is very simple. Even with marking, the learning of a player will typically be limited. This suggests that different players will likely have different views and one should be cautious when using arguments based on full cognizance or common knowledge of the game structure. Arguments based on more limited structural knowledge, like what is given in Kaneko-Kline [11], are more appropriate.

In addition, by specifying who the learner is and what are learned, our research departs significantly from the evolution/learning literature in game theory/economics (cf. Weibull [21], Fudenberg-Levine [5] and Kalai-Lehrer [9]). One negative implication from our approach is that the focus on limiting cases is no longer appropriate. It is a positive implication that our research is more related everyday memory in the psychology literature (Linton [7], [8] and Cohen [2]). Delving deeper into that literature and combining it with our simulation analysis and theoretical findings may lead to new findings about the learning and behavior of humans in game situations. Nevertheless, there is still a large distance between our study of inductive game theory together with the present simulation study and experimental psychology and we need to develop our theory and simulation study more to build bridges between those fields.

Toward building an above mentioned bridge, we would meet a lot of conceptual, theoretical and experiential problems. To explore such problems, theoretical and simulation studies as well as philosophical (methodological) discussions are required. In the following, we will mention possible extensions of our simulation study given in the present paper. But before it, would be a good idea to recall that some conceptual problems arise from our simulation study, while they have not been discussed as a theoretical problem.

An example of a new conceptual problem arising from our simulation study is the distinction between having preferences and knowing preferences. Also, we have found that the lasting time differs with information expressed in a language and an emotion such as satisfaction. To distinguish between those concepts, we need philosophical arguments as well as some experimental (psychological) study. Even this problem has questioned the treatment of information expressed in an information set, due to von Neumann-Morgenstern [22] and Kuhn [16], in the standard game theory literature (cf. Osborne-Rubinstein [20]). In Kaneko-Kline [12], the treatment of information and memory was reconsidered and reformulated. This has lead to new technical developments, like the separation of information and memory given in the concept of an information protocol in [12].

In the remaining of this section, we will mention some possible extensions of our simulation study.

Aspect 1.1: Long-term memories and decaying: We assumed that once an experience becomes a long-term memory, it will last forever. However, it would be more natural to assume that even long-term memories are subject to decay if they are not experienced again once in a while. In particular, when the regular behavior changes as in the example of Section 5.2, decay or forgetfulness about past regular behavior might become important.

This problem is related to Ebbinghous' [4] retention function which was used to describe experimental results of memory of a list of nonsense syllables. No distinction is made there between a short-term memory and a long-term memory. The retention function is typically considered as taking the shape of any curved line depicted in Figure 7.1, where the height denotes the probability of retaining a memory and it is diminishing with time[10].

It is the point relevant to our research that repetitive learning makes the probability of retention diminish more slowly. In Fig.7.1, the second solid curve is obtained when the second experience occurs while the first experience still remains as a memory. On the other hand, the dotted curve is obtained if the first experience disappeared from his memory before the the second experience. Thus, the shape of the dotted curve is the same as the first solid one. The second solid curve is flatter than the first one because of repetitive reinforcement. If the third experience occurs soon enough, we move to the third solid curve which is even flatter.

Our treatment of memory can be expressed similarly. For this, consider $(m, k) = (10, 2)$. Once the subject has an experience at t_1, he keeps it as a memory for 10 periods. In Figure 7.2, the second experience does not come to him within 10 periods, but it comes later at t_2. Then the third experience comes within 10 periods after t_2, and the memory remains forever, which we call a long-term memory in this paper.

In Ebbinghous's case, the retention function becomes flatter with more experiences, meaning that the memory has a longer expected life. A longer lived memory is more likely to be repetitively reinfoced, and so the memory may persist. Our treatment can be seen as a simplification of Ebbinghous', where we distinguish between a short-term and a long-term memory, ignoring decay.

Aspect 1.2: Intensities of experiences and preferences: We also ignored intensities of stimuli from experiences. This aspect could be important in the treatment of pref-

[10] His experiments are interpreted as implying that the retention function may be expressed as an exponential function. By careful evaluations of Ebbinghous' data, Anderson-Schooler [1] reached the conclusion that the retention function can be better approximated as a power function, i.e., the probability of retaining a memory after time t is expressed as $P = At^{-b}$.

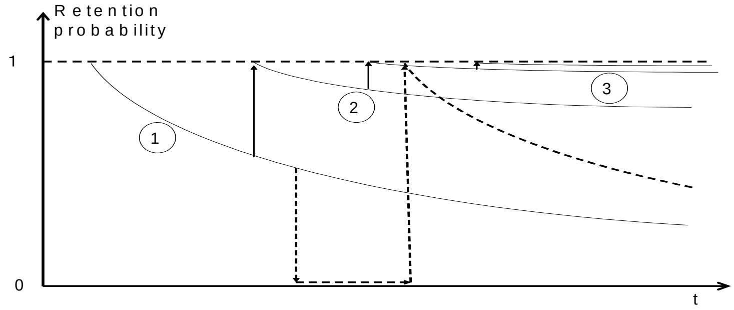

Figure 7.1: Ebbinghous' Retention Function

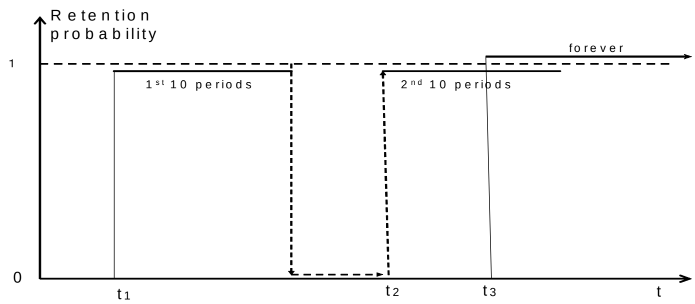

Figure 7.2: Our Retention Function

erences in Section 5. For example, only preference intensities that are beyond some threshold remain in short-term memories. The use of thresholds is similar to the need for repetition, but it could be introduced and explored in our future work.

Aspect 1.3: Two or more learners: We have concentrated our focus on the example of Mike's Bike Commuting. We are interested in learning in game situations with two or more learners (players). This has other new features like the relevant learning time. For example, I may learn over my life time, but only interact with another player for a shorter time span. In addition, how does my learning affect your learning? We might start with the "small and partial views" setting of Kaneko-Kline [11], but expect to move on as dynamics coming from communication and role playing will likely surface.

These are straightforward extensions but may expect a lot of implications to our study. We can even introduce more probabilistic factors related to decaying of long-term as well as short-term memories. However, more essential extensions are related to the consideration of internal structures of routes and inductive derivations of individual views from experiences.

Aspect 1.4: Internal Structures and subattributes: In order to simplify the analysis, we ignored the internal structures of the routes by treating each entire route as one symbolic identity. Nevertheless, inductive game theory is about the formulation of a player's beliefs about the structure of a game situation. The internal structure and subattributes are relevant to this type of analysis. In fact, the introduction of internal structures will be a key for essential developments of our simulation study as well as inductive game theory itself.

Aspect 1.5: Simulation study of inductive derivations: When Aspect 1.4 is taken into account, we should consider the connection of subattributes. This is more directly related to inductive derivations, which is originally motivated by inductive game theory. Kaneko-Kline [13] addresses this question in inductive game theory. It must be possible to study details of this inductive process by simulation study. It will give more detailed information than the theoretical study. One immediate question from this to Mike's Bike Commuting is what maps Mike can construct, and whether they are correct and complete.

Our simulation study and the aspects mentioned above provide a lot of new directions for research and implications for inductive game theory and the extant game theory.

References

[1] Anderson J. R. and L. J. Schooler, Reflections of the environment in memory, *American Psychological Society* 2, 396-408.

[2] Cohen, G., M., (1989), *Memory in the Real World*, Lawrence Erlbaum Associates Ltd. Toronto.

[3] Deutsch, D. and J. A. Deutsch, ed. (1975), *Short-term Memory*, Academic Press, New York.

[4] Ebbinghous, H., (1964, 1885), Memory: A contribution to experimental psychology, Mieola, NY: Dover Publications.

[5] Fudenberg, D., and D.K. Levine, (1998), *The Theory of Learning in Games*, MIT Press, Cambridge.

[6] Harsanyi, J. C., (1967/68), Games with Incomplete Information Played by 'Bayesian' Players, Parts I,II, and III, *Management Sciences* 14, 159-182, 320-334, and 486-502.

[7] Linton, M., (1975), Memory for Real-World Events, *Exploration in cognition*, eds. D. A. Norman & D. E. Rumelhart, Freeman Publisher, San Francisco.

[8] Linton, M., (1982), Transformations of memory in everyday life. In U, Neisser ed. Memory Observed: Rembering in natural contexts. Freeman, San Francisco.

[9] Kalai, E., and E. Lehrer, (1993), Subjective Equilibrium in Repeated Games, *Econometrica* 61, 1231-1240.

[10] Kaneko, M., (2002), Epistemic Logics and their Game Theoretical Applications: Introduction. *Economic Theory* 19, 7-62.

[11] Kaneko, M., and J. J. Kline, (2007a), Small and Partial Views derived from Limited Experiences, University of Tsukuba, SSM.DP.1166, University of Tsukuba.

http://www.sk.tsukuba.ac.jp/SSM/libraries/pdf1151/1166.pdf.

[12] Kaneko, M., and J. J. Kline, (2007b), Information Protocols and Extensive Games in Inductive Game Theory, Bond University, Business Paper MS #1005, http://epublications.bond.edu.au/business_pubs/6

[13] Kaneko, M., and J.J. Kline (2008), Various Types of Induction and Constructed Views (mimeo).

[14] Kaneko, M., and A. Matsui, (1999), Inductive Game Theory: Discrimination and Prejudices, *Journal of Public Economic Theory* 1, 101-137. Errata: the same journal 3 (2001), 347.

[15] Kleene, S. C., (1967), *Mathemtical Logic*, John Wiley & Sons Inc., New York.

[16] Kuhn, H. W., (1953), Extensive Games and the Problem of Information, *Contributions to the Theory of Games II*, Kuhn, H. W. and A. W. Tucker, eds. 193-216. Princeton University Press.

[17] Luce, R. D., and H. Raiffa (1957), *Games and Decisions*, John Wiley and Sons Inc., Boston.

[18] Malinvaud, E., (1972), *Lectures on Microeconomic Theory*, North-Hollond. Amsterdam.

[19] Tulving, E., (1983), *Elements of Episodic Memory*, Oxford University Press, London.

[20] Osborne, M., and A. Rubinstein, (1994), *A Course in Game Theory*, MIT Press, Cambridge.

[21] Weibull, J. W., (1995), Evolutionary Game Theory, MIT Press. London.

[22] von Neumann, J., and O. Morgenstern, (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.