Inductive Game Theory: A Simulation Study of Learning a Social Situation*

Eizo Akiyama, Ryuichiro Ishikawa, Mamoru Kaneko and J. Jude Kline

Additional information is available at the end of the chapter

1. Introduction

Inductive game theory (IGT) aims to explore sources of beliefs of a person in his individual experiences from behaving in a social situation. It has various steps, each of which already involves a lot of different aspects. A scenario for IGT was spelled out in Kaneko-Kline [15]. So far, IGT has been studied chiefly in theoretical manners, while some other papers targeted applications and conducted an experimental study. In this chapter, we undertake a simulation study of a player's learning about some details of a social situation. First, we give a brief overview of IGT, and its differences from the extant game theories. Then, we explain several points pertinent to our simulation model.

1.1. Developments of inductive game theory

The scenario for IGT given in [15] consists of three main stages:

(1) Experimental Stage: making trials-errors and accumulating experiences;

(2) Inductive Derivation Stage: construction of an individual view from accumulated experiences;

(3) Analysis/Use Stage: uses of the derived view for behavioral revision.

Fig.1 describes the relationships among the three stages[1]. The process starts with the experimental stage, where a player makes trials-errors, and accumulates memories from experiences. In the second stage, each player constructs an individual view from accumulated experiences, which is based on induction; this is the reason for the title "inductive" game

---

[1] We may regard (1) and (2)-(3), respectively, as corresponding to the *experiencing self* and the *remembering self* in Kahneman [11, p.381]. Kahneman talks about various examples and aspects relevant to this distinction.
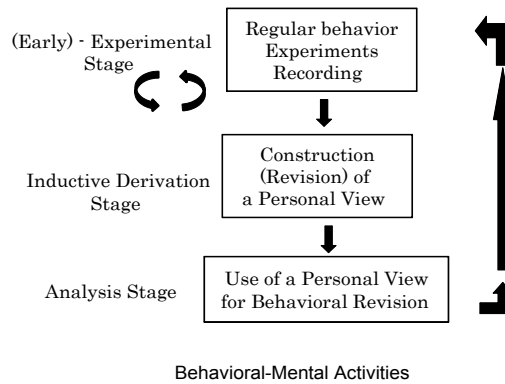

INTECH
open science | open minds

Behavioral-Mental Activities

Figure 1. Three Stages of IGT

theory. In the third stage, once a player has built his view, he uses it for his decision making or behavioral revision. After the third stage, the process goes to the first stage, and those stages may cycle.

Each stage already includes a lot of new problems. To study those problems, we borrow concepts from the extant game theories, but often we need to think about whether some can or cannot be used for IGT and whether to modify them for IGT, since they often rely upon the presumptions of the extant game theories.

In Kaneko-Matsui [19] and Kaneko-Kline [15], [16], [17], we have focused on the second and third stages. The first stage of making trials-errors and accumulating memories was discussed, but described in the form of informal postulates. Taking the resulting sets of accumulated memories from trials and errors as given, the second and third stages are formulated in a theoretical manner. However, the first stage is of a very different nature from the other two, and each player's bounded cognitive ability is crucial. For this, we may take two approaches: experimental and simulation. Takeuchi *et al.* [22] conducted an experimental study, and here, we take a simulation method.

It would be helpful to discuss, before giving a description of our simulation study of IGT, how IGT differs from two main stream approaches in the recent game theory literature: the classical *ex ante* decision approach and the evolutionary/learning approach. The contrasts between them will 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 [8] for the incomplete information game and Kaneko [13] for the epistemic logic approach to decision making in a game). In this approach, the beliefs/knowledge is given *a priori* without asking their sources. Thus, IGT is relevant for exploring sources of beliefs and knowledge in experiences.

Contrary to this, the evolutionary/learning approach (cf., Weibull [24], Fudenberg-Levine [6], and Kalai-Lehrer [12]) targets "learning". However, this approach does not ask the question of the emergence of beliefs/knowledge; instead, their concern is typically the convergence of the distribution of actions to some equilibrium. 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 learning of the structure or details of the game; typically it is not specified who the learner is and what is learned. When we work on an individual's learning, we should make these questions explicit.

If the learner is an ordinary person, the convergence of behavior in the limit is not very relevant to his learning. Finiteness of life and learning must be crucial. Here, "finite" is "shallowly finite", rather than the negation of infinity in mathematics. Consequently, we conduct simulations over finite spans of time corresponding to the learning span of a single human player. Our simulation indicates various specific components affecting one's finite learning, while they are not relevant in the limiting behavior.

1.2. Simulation study of a social situation

Now, we discuss several important points of our simulation model.

(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 only as a small part of the social world for that person. We choose a simple and casual example called "Mike's Bike Commuting". In this example, the learner is Mike, and he learns the various routes to his work. Using this example, the time span and the number of reasonable repetitions for the experiment become explicit.

We study a one-person problem, but it should not be regarded as isolated from society. It is a small part of Mike's social world.

(2): *Ignorance of the situation:* At the beginning, Mike has no prior beliefs/knowledge about the town. His colleague gave a coarse map of possible alternative routes without precise details, and suggested one specific route from his apartment to the office. Mike can learn the details of these routes only if he experiences them. We question how many routes Mike is expected to learn after specific lengths of time.

(3): *Regular route and occasional deviations:* Mike usually follows the suggested route, which we call the regular route. Occasionally, when the mood hits him, he takes a different route. This is based on the basic assumption that his energy/time to explore other routes is scarce. Commuting is only a small part of his social world, and he cannot spend his energy/time exclusively for exploring those routes.

(4): *Short-term and long-term memories:* We distinguish two types of memories for Mike: short-term and long-term. Short-term memories form a finite time series consisting of past experiences, and they will be kept only for some finite length of time, perhaps a few days or weeks; after then they will vanish. However, when an experience occurs with a certain frequency, it becomes a long-term memory. Long-term memories are lasting.

In our theory, the transition from a short-term to a long-term memory requires some repetition of the same experience within a given period of time. This is based on the general idea that memory is reinforced by repetition. Our formulation can be regarded as a simplified version of Ebbinghous' [5] retention function.

(5): *Finiteness and complexity:* Our learning process is formulated as a stochastic process. Unlike other learning models, we are not interested in the convergence or limiting argument. As stated above, the time structure and span are finite and short. In our example, we discuss how many times Mike has experienced a particular route after a half year, one year, or ten

years. We will find many details, which are highly complex even in this simple example. We analyze those details and find the lasting features in Mike's mind.

(6): *Marking salient choices as important:* Although the situation is extremely simple, it is difficult for Mike to fully learn the details of the entire town even after several years. We consider the positive effect on learning by "marking", introduced in Kaneko-Kline [14]. If Mike marks some "salient" choice as "important", and restricts his trial-deviations to the marked choices, then we find that his learning is drastically improved. Imperfections in a player's memory make marking important for learning. Without marking, experiences are infrequent and lapse with time. Consequently, his view obtained from his long-term experiences could be poor and small. By marking, he focuses his attention on fewer choices, and successfully retains more as long-term memories.

Up to here, we study how many times Mike needs to commute in order to learn some routes. Precise objects Mike possibly learns are not targeted. There are two directions of departure from this study. One possibility is to study Mike's learning of internal components of routes, and the other is about relationships between routes. Of course, to study both in an interactive way is possible. In this paper, however, we consider a problem of the latter sort, namely, Mike's learning of his own preferences from experiences.

(7): *Learning preferences:* Here, we face new conceptual problems. We should make a distinction between having preferences and knowing them. We assume that Mike has well-defined complete preferences, but his knowledge is constrained to only some part by his experiences. Also, it is important to notice that learning one's preferences differs from keeping a piece of information. 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 of outcomes close in time. Consequently, marking alternatives becomes even more important for obtaining a better understanding of his own preferences.

In our simulation study up to Section 4, we will get some understanding of relevant "shallowly finite" time spans for ordinary life learning. Our study on learning preferences in Section 5 is more substantive than the studies up to Section 4. However, we will not go to the direction to a study of learning of internal structures of routes. This will be briefly discussed in Section 7.

The chapter 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. In Section 4, we introduce the concept of "marking", and observe its positive effects on learning. In Section 5, we consider the problem of learning his preferences. In Section 6, we carry out a sensitivity analysis of changing various parameters describing Mike's learning and memory characteristics. Section 7 is devoted to a discussion our results and their implications for IGT 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 and indicates one route shown by the dotted line. Mike starts commuting every morning and evening, five days a week, that is, 10 times a week. From the beginning, he wants to know the details of those
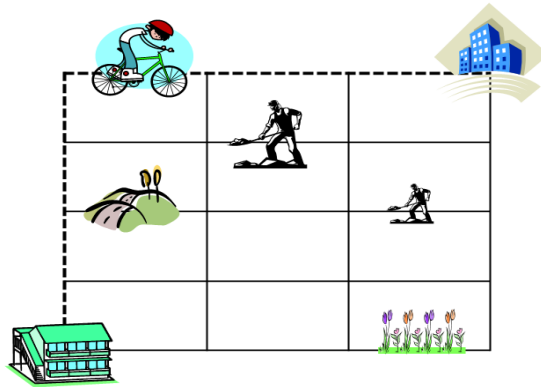
Figure 2. A Map of the Town

routes, but the map is simple and coarse. He decides to explore some alternative routes when the mood hits him, but typically he is too busy or tired and resorts to the *regular route* suggested by the colleague[2].

The town has a lattice structure: His apartment and office are located at the south-west and north-east corners. To have a route of the shortest distance from his apartment to the office, he should choose "North" or "East" at each lattice point; such a route is called a *direct* route. There are 35 direct 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 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.3. He himself may not be conscious of these probabilities or of this tree. In sum, on average, he makes a deviation twice a week to any of the other routes with equal probability.

After following route a_l, he gets some impressions and understanding of a_l. In this paper we do not study the details of a_l that he learns; instead, we study conditions for an experience to remain in his mind as a long term memory.

As mentioned in Section 1, 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 in m trips, experience a_l becomes a long-term memory. Long-term memories form a set of experiences without time-structure or frequency[3].

---

[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 IGT given in Kaneko-Kline [15]. However, this makes our simulation study much more complicated. We will keep our study as simple as possible.

[3] This lack of time structure and frequency is motivated by bounded rationality of the player. Limitations on his memory and computation abilities lead him to ignore some aspects like the time structure and frequency of long term memories.

In our simulation, 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.4, 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.

We consider another parameter T, denoting the total number of trips (time span). For example:

$$\text{after a half year, } T = 2 \times 5 \text{ (days)} \times 25 \text{ (weeks)} = 250;$$
$$\text{after 1 year, } T = 2 \times 5 \text{ (days)} \times 50 \text{ (weeks)} = 500;$$
$$\text{after 10 years, } T = 2 \times 5 \text{ (days)} \times 500 \text{ (weeks)} = 5000.$$

Our simulation will be done by focusing on the half year and 10 year time spans. In Mike's Bike Commuting, the number of available routes is 35, but later, this 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*:

$$F = [s, p; (m, k)]. \tag{1}$$

We always assume that in the case of a deviation, a route other than a_0 is chosen with equal probability $1/(s - 1)$.
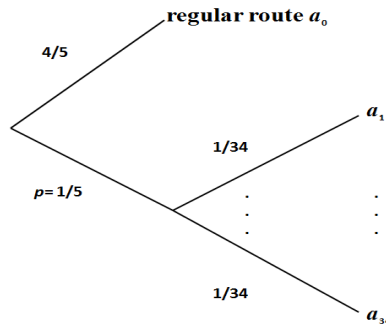

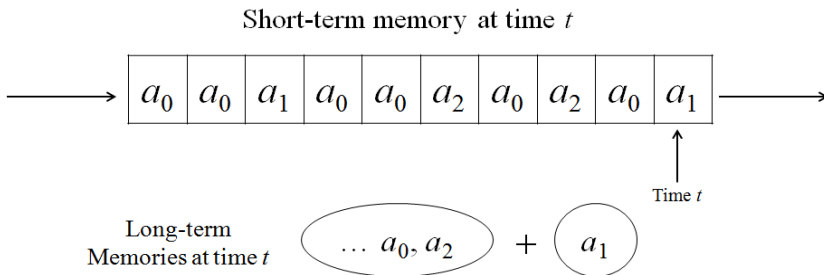**Figure 3.** Decision Tree of each Trip of Commuting

Figure 4. Short-Term and Long-Term Memories

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

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

$$\text{the probability that all routes become long-term memories} \tag{2}$$
$$\text{tends to 1 as } T \text{ tends to infinity.}$$

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. Our interest, however, is in finite learning. Our findings by simulation for the finite learning periods of $T = 250$ and $T = 5000$ differ significantly from the convergence result. This suggests that focusing on convergence results does not inform us about finite learning.

3. Preliminary simulations and the method of simulations

We start in Section 3.1 by giving simulation results for the case of $s = 35$. The results show 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 give a brief explanation of our simulation method and the meaning of "probability".

3.1. Simulation results for $s = 35$

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

The row for a_0 shows that the probability of the regular route a_0 being a long-term memory is already 1 at $T = 250$ (a half year). This "1" is still an approximation result meaning it is very close to 1.

The row for a_l ($l \neq 0$) is more interesting. The probability that a specific a_l 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.

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

Table 1

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

Table 2

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

Table 3

Some reader may have expected that the probability for $T = 250$ 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 sequences. Hence, the probability turns out not to be negligible[4]. 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[5], and thus we regard the limiting convergence result (2) as only a reference.

The cases of $T = 250$ and 5000 are 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 2, which describes the probability of exactly r routes (the regular route and $r - 1$ alternative routes) being long-term memories in 35 routes:

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. For $r = 3$, one route must be regular, but the other 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 very small. This means that although Mike learns about 2 alternative routes, it is hard to predict with much accuracy which pair would be learned.

At $T = 5000$, i.e., ten years later, Mike's learning is described by Table 3.

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. Because most of the distribution lies between 25 and 29 routes, we find that there are 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. This probability can be obtained from the probability 0.765 in Table 1 by the equation:

$$\binom{34}{24} \times (0.765)^{24} \times (1 - 0.765)^9 \doteq 0.109.$$

---

[4] 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 not to have a pair of students of the same birthday. However, the exact calculation tells that the probability is about 0.97.

[5] Our model without decay of long-term memories is likely to be inappropriate for 56 years.
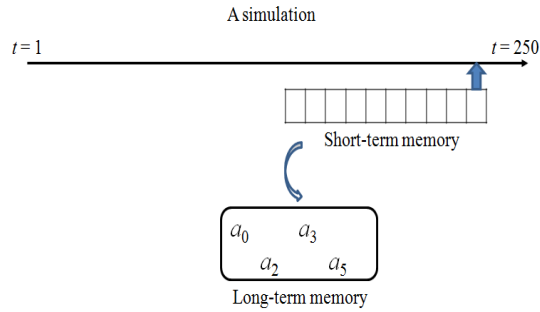
Figure 5. A Simulation up to $T = 250$

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 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.

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.

After all, the above calculations indicate that "finiteness" involved in our ordinary life is far from "large finiteness" appearing in the convergence argument in mathematics. In this sense, we are facing shallowly finite problems, which was emphasized in Section 1. In Sections 4 and 5, we will discuss related problems to this issue from different perspectives.

3.2. Simulation method

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 $T = 30$, but beyond $T = 40$, it is practically impossible in the sense that for $T = 50$, it takes decades to calculate with current (year 2007) computers using our analytical method. This is caused by the limited length of short-term memory and multiple occurrences needed for a long-term memory.

We take the relative frequency of a given event over many simulation runs instead of computing probabilities analytically. 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.3, 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.5. One simulation run gives a set of long-term memories: In Fig.5, 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:

$$\frac{\#\{\text{simulation runs with } a_l \text{ as a long-term memory}\}}{100,000} \qquad (3)$$

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.

We compare some results from simulation with the results obtained 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 an accuracy of 10^{-4}.

The robustness of the frequency (probability) 0.069 in Table 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, $68,596,187$), both of which translate to the frequency 0.069 when rounding off to three decimal places.

In sum, we calculate the "probability" of an event as the relative frequency over numerous simulation runs since the analytic calculation is difficult for the large finite time spans and simulation frames under consideration.

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

We now show how "marking", introduced in Kaneko-Kline [14], 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 repeating an experience rises by reducing the number of alternative routes. In Section 4.1, we consider the case where Mike marks only four alternative routes in addition to the regular one. We see a dramatic increase in his learning of alternative routes. In Section 4.2, we show how a more planned approach can improve the effect of "marking" on his learning.

4.1. Marking five salient routes and simulation results

Suppose that Mike decides to mark some routes from his map for his exploration. He uses two criteria:

(i) He chooses routes having a scenic hill or flowers;

(ii) He avoids construction sites.

Then, he marks only four alternative routes, which are depicted in Fig.6. Adding the regular route a_0, 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 simulation frame to $F = [s, p : (m, k)] = [5, 1/5 : (10, 2)]$ for $T = 250$ or 5000. The probability of a_l $(l \neq 0)$ being a long-term memory is calculated by our simulation method and is given in Table 4:
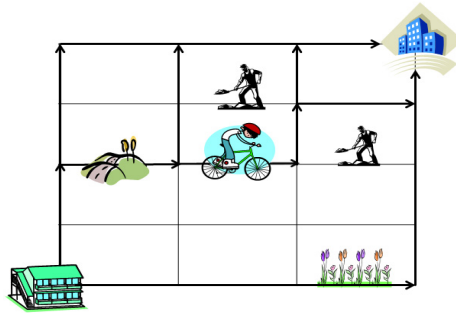
Figure 6. Five Marked Routes

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

Table 4

Table 5 lists the length of time needed to obtain the probability 0.99 that an alternative route a_l ($l \neq 0$) is a long-term memory. With marking he needs only 425 trips (10.2 months), as opposed to the 28, 253 trips (more than 56 years) without marking.

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

Table 5

We also have calculated, and presented in Table 6, the probability that exactly r ($= 1, 2, 3, 4, 5$) routes are long-term memories at $T = 250$. The average number of routes learned is 4.9. Table 7 states that the average time for Mike to learn all 35 routes 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.

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

Table 6

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

Table 7

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).

The probability of that he finishes his entire exploration in 3.5 years is $(0.886)^7 \doteq 0.427$, 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 a rather small number of unlearned routes to remain; the event of 3 remaining is rare event occurring with only probability 0.03. With high probability, Mike's learning finishes within 4 years.

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

Here, we consider Mike's learning of his own preferences. Mike finds his own preferences based on comparisons between experienced routes. First, we 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 important. 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

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. By recalling the original meaning of "preferences", however, we can connect them with experiences. Since an experience of each route gives some level of satisfaction, comparisons between satisfaction levels can be regarded as his preferences. Here, preferences are assumed to be 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.

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

Table 8

We assume that Mike's inherent preference relation over the routes is complete and transitive. A preference between two routes is experienced only by comparing the two satisfaction levels from those routes[6][7]. A feeling of satisfaction typically emerges in the mind (brain) without tangible pieces of information. Such a feeling 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 route can be compared with that of another only if these have happened closely in time.

We formulate a preference comparison between two routes as an experience. 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*. Our present target is second-order experiences.

Consider Mike's learning of such second-order experiences in the simulation frame $F = [s, p : (m, s)] = [5, 1/5 : (10, 2)]$ with $T = 250$ or 5000. A short-term memory is now treated as a sequence of length 10. Consecutive routes can be compared to form preferences over pairs. For example, in Fig.7, the short-term memory is the sequence of 10 pairs $\langle a_1, a_0 \rangle, \langle a_0, a_0 \rangle, ..., \langle a_3, a_0 \rangle$. We treat them as unordered pairs, e.g., the pairs $\langle a_1, a_0 \rangle$ and $\langle a_0, a_1 \rangle$ in $t - 9$ and $t - 5$ are treated as the same. These second-order experiences may become long-term memories.

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.7, $\langle a_0, a_1 \rangle$ occurred twice, and hence it becomes a long-term memory. We require these consecutive unordered pairs be disjoint; for example, (a_0, a_3) and (a_3, a_0) occurred twice having the intersection a_3, so these occurrences are not counted as two.
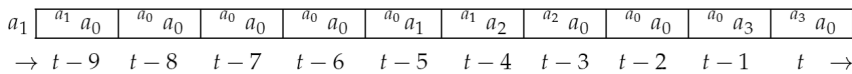
a_1	a_1 a_0	a_0 a_0	a_0 a_0	a_0 a_0	a_0 a_1	a_1 a_2	a_2 a_0	a_0 a_0	a_0 a_3	a_3 a_0
$\rightarrow$	$t-9$	$t-8$	$t-7$	$t-6$	$t-5$	$t-4$	$t-3$	$t-2$	$t-1$	t $\rightarrow$

Figure 7

---

[6] This should be distinguished from the notion of "revealed preferences" (cf. Malinvoud [20]) where 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.

[7] Our problem is how a person learns his own preferences from experiences, but not how his preferences emerge. In this sense, our problem is not "endogenous preferences". Nevertheless, our problem includes partial and/or false understanding of one's own preferences; thus, it is potentially related to the field of endogenous preferences. See Bowles [2] and Ostrom [21] for the literature on endogenous preferences, and see also Kahneman [11] for other aspects related to this literature as well as our problem.

r	4	5	6	7	8	9	10
10 years later	1.07×10^{-3}	0.0155	0.079	0.215	0.329	0.269	0.0913
20 years later	1.59×10^{-15}	7.86×10^{-5}	0.0016	0.0179	0.111	0.366	0.504

Table 9. Probabilities of preference learning after 10 and 20 years

The computation result is given in Table 8 with $l, l' = 1, 2, 3, 4$ and $l \neq l'$. 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 only about 2 years, the probability is already 1[8].

We find in the right column of Table 8 that Mike's 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 disjointedly. This means that his learned preferences are very incomplete even after quite some time.

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, which is depicted as the left diagram of Fig.8. 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 illustrated in the middle diagram of Fig.8. 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.

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 9 lists the probabilities that exactly r of his preferences are learned. Recall that there are $\binom{5}{2} = 10$ comparisons. Even after 10 years, Mike is still learning his own preferences over alternative routes. After 20 years, however, he learns much more about his preferences. As it happens, by the time Mike is able to get to taste the rough with the smooth, he is already old.
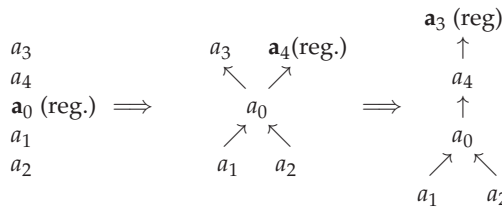

Figure 8

---

[8] One might wonder why the value of 0.981 for a comparison between a_0 and a_l is higher 0.970 for just learning a route a_l in Table 4. This can be explained by the counting of pairs at the boundary. For example, the comparison between a_0 and a_1 appearing in Table 8 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.

5.2. Maximizing preferences

The results of the previous subsection tell us that it is 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 he finds a better route a_l than a_0 in those comparisons, then he chooses a_l (arbitrarily, if there are multiple) 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.

The final result of this process 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 Fig.8; he has a strict preference ordering $a_3 \succ a_4 \succ a_0 \succ a_1 \succ a_2$, where a_0 is the regular route. After some time, he learns his preferences described in the middle diagram. In this case, it is very likely that only his preferences between a_0 vs. a_l ($l \neq 0$) are learned. The arrow $\rightarrow$ indicates the learned preferences.

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 part. The transition from the left column to the middle one in Fig.8 needs the average time 136.2 (3.3 months). When he reaches the middle diagram, he stores the preferences over a_0, a_1 and a_2.

In the middle diagram of Fig.8, he starts comparing between a_3 and a_4. Here, a_4 is taken as the new regular route. 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 rightmost diagram takes the average time of $136.2 + 11.0 = 147.2$ trips (3.5 months).

We have $5! = 120$ possible preference orderings over a_0, a_1, a_2, a_3, a_4 and a_5. We classify them into 5 classes by the position of a_0. Here we consider only the other two cases: a_0 is the top or the bottom. When a_0 is the top, only one round of comparing a_0 to other a_l is enough to learn that a_0 is his most preferred route. This takes the average time 136.2 (3.3 months), which is the same as the time for the transition to the middle of Fig.8. In the case with the top a_0, however, Mike learns no other preferences.

Consider the case where 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
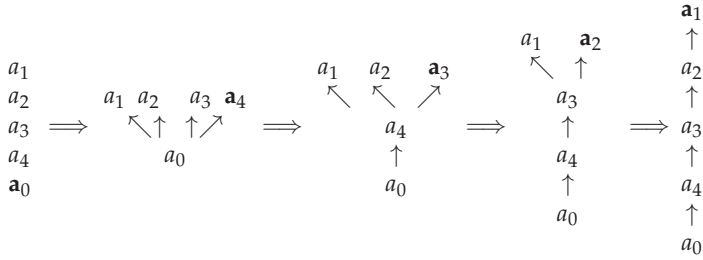
$$
\begin{array}{c}
a_1 \\
a_2 \\
a_3 \implies \\
a_4 \\
\mathbf{a_0}
\end{array}
\quad
\begin{array}{c}
a_1 \; a_2 \;\; a_3 \; \mathbf{a_4} \\
\nwarrow \uparrow \;\; \uparrow \nearrow \\
a_0
\end{array}
\implies
\begin{array}{c}
a_1 \;\; a_2 \;\; \mathbf{a_3} \\
\nwarrow \;\; \nwarrow \;\; \nearrow \\
a_4 \\
\uparrow \\
a_0
\end{array}
\implies
\begin{array}{c}
a_1 \quad \mathbf{a_2} \\
\nwarrow \;\; \uparrow \\
a_3 \\
\uparrow \\
a_4 \\
\uparrow \\
a_0
\end{array}
\implies
\begin{array}{c}
\mathbf{a_1} \\
\uparrow \\
a_2 \\
\uparrow \\
a_3 \\
\uparrow \\
a_4 \\
\uparrow \\
a_0
\end{array}
$$

Figure 9. Transitions with learning preferences

for finding the top needs 4 transitions. Fig.9 depicts the slowest case: The total average time is $136.2 + 78.0 + 36.4 + 11.0 = 261.6$ (6.3 months); the bold letter means the regular route. By this process he finds his 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. On the other hand, if he finds the top slowly, he would have a much richer knowledge of his own preferences[9].

6. Sensitivities with parameter changes

We have seen the effects of changes of s and T on Mike's learning determined by the simulation frame $F = [s, p; (m, k)]$. In this section, we briefly consider the sensitivity of the simulation results 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 parameters (m, k) are of a different nature. First, we keep in mind that our intention is to capture casual everyday learning. While p is regarded as externally given, it may be controlled by Mike in an effort to learn more about alternative routes. The parameters m and k may also be within Mike's control, but because they describe his memory ability, changing them may require greater effort on his part than increasing p. 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 10 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.

On the other hand, Table 11 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 12 and 13.

---

[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-order 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.

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

Table 10. $s = 5$ and $m = 10$

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

Table 11. $s = 5$ and $k = 2$

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

Table 12. $s = 5$

(m, k)	$T = 250$	$T = 5000$
$(10, 2)$	0.069	0.765
$(20, 3)$	0.007	0.140

Table 13. $s = 35$

$p \setminus T$	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 14. $s = 5$

Table 13 shows that increasing both k and m implies that Mike's 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 12, 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 14 and 15:

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 rightmost columns of Tables 14 and 15 also list the average number

$p \setminus T$	250	5000	Av.no
0.05	0.005	0.091	215707
0.1	0.018	0.312	54893
0.2	0.069	0.765	14223
0.3	0.143	0.957	6548.5

Table 15. $s = 35$

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

The example of Mike's bike commuting is a small everyday situation and provides insights to our everyday behavior. It is designed to capture several aspects of a human behavior in a social world. One important aspect is that the life span of a human being has a definite upper bound. Mike's bike commuting is used to compute what learning is possible within his life span. Also, our target situation is partial relative one person's entire social world. In this respect, the regular behavior is a consequence of time/energy saving and infrequent deviations are exploration behavior. We conducted various simulations to see effects of those aspects.

Consider some implications of our simulation study to related literatures. Our original motivation was, from the viewpoint of IGT, 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 person to learn the full structure of a game, unless it is very simple. Even with marking, the learning will typically be limited. A focus on limiting cases is no longer appropriate. This leads us to deviate entirely from the literature of evolutionary/learning approach mentioned in Section 1.1.

Our research is more related to everyday memory in the psychology literature (Linton [9], [10] and Cohen [3]). Yet, there is a large distance between our study and experimental psychology. To build a bridge between those fields, we need to develop our theory as well as experimental and simulation studies. Kaneko-Kline [14] had a theoretical study in this direction by introducing a measure of the size of an inductively derived view and considering the effects of marking. This is one direction among many other possible extensions.

In the following, we mention several other possible extensions.

Aspect 1: Long-term memories and decaying: We assume 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 unless they are kept experienced once in a while. In
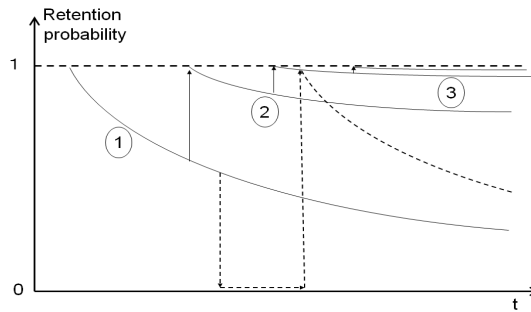
Figure 10. Ebbinghous' Retention Function

particular, when the regular behavior changes as in Section 5, decay or forgetfulness about past regular behavior might become important. This is relevant to the problem of Section 4.2.

The above problem is related to Ebbinghous' [5] retention function which was used to describe experimental results of memory of a list of meaningless syllables. There, no distinction is made between a short-term memory and a long-term memory. The retention function is typically considered as taking the shape of a curved line depicted in Fig.10, where the height denotes the probability of retaining a memory and it is diminishing with time[10].

It is more relevant to our research that repetitive learning makes the probability of retention diminish more slowly. In Fig.10, 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 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 Fig.11, 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.

In Ebbinghous' 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 reinforced, and so the memory may persist. Our treatment can be seen as a simplification of Ebbinghous' retention function, where we distinguish between a short-term and a long-term memory without decay.

This direction may become even more fruitful with an experimental study.

Aspect 2: Intensities of experiences and preferences: We also ignored intensities of stimuli from

---

[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 $P = At^{-b}$.
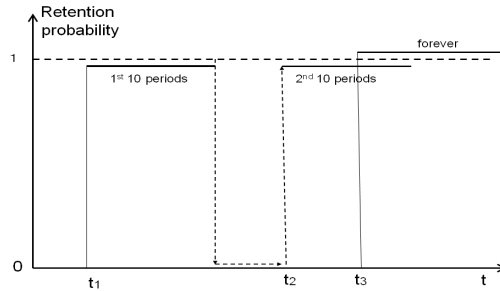
Figure 11. Our Retention Function

experiences. This aspect could be important in the treatment of preferences 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. The concept of "marking" (saliency) is closely related to this problem. It is a topic for future work.

Aspect 3: Two or more learners: We have concentrated our focus on the example of Mike's bike commuting. Our original interests are in learning in game situations with two or more learners (persons)[11]. This has other new features: For example, how does his learning affect the other's learning? In particular, when we consider the other person's understanding, possibly by switching social roles, it affects the persons' behaviors drastically, e.g. emergence of cooperation may be observed. These possibilities are studied in Kaneko-Kline [18]. In that setting, the domain of experiences plays essential roles, for which a simulation study must be informative[12].

These extensions may generate a lot of implications for IGT. 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 4: Internal Structures and subattributes: We ignored the internal structure and subattributes of each route in the town by treating it as one entity. Nevertheless, IGT is about the formation of a person'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 such internal structures will be a key for essential developments of our simulation study as well as IGT itself.

When this is taken into account, an inductive derivation may be regarded as drawing a picture by connecting one subattribute with another. This is originally motivated in Kaneko-Kline

---

[11] Hanaki *et al.* [7] studied the convergence of behaviors in a 2-person game, where each player's learning of payoffs is formulated in the way of the present paper but his behavior is formulated as a mechanical statistical process following the learning literature. Then, they studied behavior of outcomes in life spans of middle range. Their approach did not take purely the viewpoint of IGT in that a player consciously makes a behavior revision once he has a better understanding of a game situation. Nevertheless, it would give some hint to our further research on IGT.

[12] These aspects are considered in an experimental context in Takeuchi, *et al.* [22], but are not connected to a simulation study.

[15]. Such a process is partially discussed in a theoretical manner in Kaneko-Kline [17]. However, a simulation study will give more detailed information. One immediate question from this to Mike's bike commuting: When Mike is told only one route from the colleague without a map of the town, what kind of a map can Mike construct? After given periods of time, how correct and complete is it?

In sum, simulation studies of those new aspects provide implications for IGT and a lot of new directions for research.

Author details

Eizo Akiyama, Ryuichiro Ishikawa, Mamoru Kaneko
Faculty of Engineering, Information, and Systems, University of Tsukuba, Japan

J. Jude Kline
School of Economics, University of Queensland, Australia

8. References

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

[2] Bowles, S. (1998), Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions, *Journal of Economic Literature*, 36(1), 75-111.

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

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

[5] Ebbinghous, H., (1964, 1885), *Memory: A contribution to experimental psychology*, Dover Publications, New York.

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

[7] Hanaki, N., R. Ishikawa, and E. Akiyama, (2009), Learning Games, *Journal Economic Dynamics & Control* 33, 1739-1756.

[8] 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.

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

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

[11] Kahneman, D. (2011), *Thinking, Fast and Slow*, Penguin, London.

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

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

[14] Kaneko, M., and J. J. Kline, (2007), Small and Partial Views derived from Limited Experiences, SSM. DP. No.1166, University of Tsukuba. http://www.sk.tsukuba.ac.jp/SSM/libraries/pdf1151/1166.pdf

[15] Kaneko, M., and J. J. Kline, (2008a), Inductive Game Theory: A Basic Scenario, *Journal of Mathematical Economics* 44, 1332-1363. Corrigendum: the same journal, 46, pp.620-622, 2010.

[16] Kaneko, M., and J. J. Kline, (2008b), Information protocols and extensive games in inductive game theory, *Game Theory and Applications* 13, 57-83.

[17] Kaneko, M., and J. J. Kline, (2009a), Partial Memories, Inductively Derived Views, and their Interactions with Behavior, to appear in *Economic Theory*. DOI: 10.1007/s00199-010-0519-0

[18] Kaneko, M., and J. J. Kline, (2009b), Transpersonal Understanding through Social Roles, and Emergence of Cooperation. DP. No.1228, University of Tsukuba. http://www.sk.tsukuba.ac.jp/SSM/libraries/pdf1226/1228.pdf

[19] 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, p.347, 2001.

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

[21] Ostrom, E. (2000), Collective Action and the Evolution of Social Norms, *Journal of Economic Perspective*, 14(3), 137-158.

[22] Takeuchi, A., Y. Funaki, M. Kaneko, and J. J. Kline, (2011), An Experimental Study of Behavior and Cognition from the Perspective of Inductive Game Theory, DP. No.1267, University of Tsukuba. http://www.sk.tsukuba.ac.jp/SSM/libraries/pdf1251/1267.pdf

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

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