Two Dialogues on Epistemic Logics and Inductive Game Theory[*,†]

Mamoru Kaneko[‡] and J. Jude Kline[§]

February 2002; and September 2010

Abstract

These two dialogues are between two professional people on a new field called "epistemic logic and inductive game theory". At the time of the first dialogue, one speaker is already a specialist and has been working in this field for a long time. The other is a game theorist, who is both younger and a novice in the field. Dialogue I takes place in January 2002: They start discussing the Konnyaku Mondô and find that it has many implications for the foundation and scope of game theory. Dialogue II occurs 8 years later following their development of inductive game theory. Now, they step back to recall what they did, as well as their trials and failures during those years. Moving forward, they discuss future research including a bridge between inductive game theory and epistemic logic.

Key Words: Epistemic logic, Common knowledge, Mutual misunderstanding, Inductive game theory, Experience, Inductive derivation, Behavioral use of a view

Dialogue I: 2002

[Setting: In January 2002, Jan Hammer, a lecturer from a foreign land, is visiting a well known institution in a remote research village north-east of Tokyo. Kurai Shinzuki, a prominent professor at the institute, has agreed to engage in discussions with the visitor. Part I takes place in a laboratory with dim lighting and inadequate heating.]

---

[*]The authors are partially supported by Grant-in-Aids for Scientific Research No.21243016, Ministry of Education, Science and Culture.

[†]The authors thank Lisa Kline and Oliver Schulte for comments on earlier versions of this paper.

[‡]Institute of Policy and Planning Sciences, University of Tsukuba, Ibaraki 305-8573, Japan (kaneko@shako.sk.tsukuba.ac.jp)

[§]School of Economics, University of Queensland, Brisbane, QLD 4072, Australia (j.kline@uq.edu.au)

1. Konnyaku Mondô

Jan: Hey, Shinzuki, I heard from your graduate students that you had discussed a Japanese comic story, called something like the Cognac Melon. They claimed it is quite interesting and important for game theory. Can you tell me that story?

Shinzuki: Ha, ha. I think it must be the Konnyaku Mondô, instead of the Cognac Melon. I can, of course, explain the story to you. But didn't my graduate students give an adequate explanation?

Jan: No, they didn't. They seemed unable to explain it since they didn't read the story, but only heard it from you. One of them claimed it was written in classical Japanese which was only understandable by some old professors. Was it written so long ago?

Shinzuki: No, no, it was written from an oral presentation only about 100 years ago.[1] It must be difficult for young people; in fact, I confess I feel it was slightly difficult for me too.

Jan: Huhm..., Japanese has changed a lot in only 100 years, hasn't it? Anyway, let's have the Cognac Melon..., sorry, Konnyaku Mondô.

[Shinzuki assumes a more professorial mode of speaking.]

Shinzuki: The literal translation of the Konnyaku Mondô is "Devil's Tongue Jelly Dialogue."

Jan: What? What is "devil's tongue jelly"?

Shinzuki: It is a food product, like a jelly, made from roots of the Konnyaku plant. You had it at the sake restaurant we went to last week. It is usually brown and looks like a devil's tongue.

Jan: Yes, yes, I remember it. But, why is that funny jelly food important to game theory?

Shinzuki: The jelly food itself is not relevant to game theory, of course, but the story of the Konnyaku Mondô is related to game theory in many respects. It is about false beliefs and the subjective nature of people's thoughts. It is an example of how inconsistent subjective thoughts may evolve, even with common knowledge and communication.

Jan: Is that true? It sounds incorrect, or if it is true, it has serious implications for game theory. Please explain the story to me.

Shinzuki: I'm not sure I will be able to reconstruct the entire story in a manner meaningful to you. Nonetheless, I shall make an attempt.

[Shinzuki now assumes an even more professorial, but somewhat dramatic, tone of

[1] pp.61–70 in [28]. The Konnyaku Mondô is from the *rakugo* form of traditional Japanese comic storytelling. In this tradition, a story is performed by a single actor.

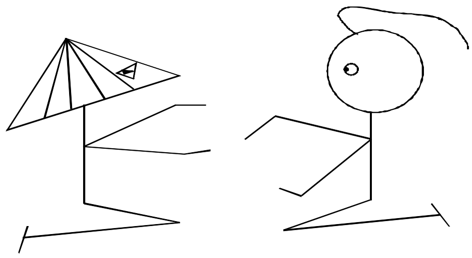

Figure 1.1: Konnyaky Mondô by Monk and Rokubei

a storyteller.]

"There was a temple without a monk for some period. A Konnyaku maker, named Rokubei, who had lived next to the temple, moved into the temple and started pretending to be a monk. One day, a Zen Buddhist monk visited the temple and challenged the new master of the temple to a dialogue on Buddhist thoughts. Since Rokubei had no idea about Buddhist dialogue, he refused at first but eventually agreed."

Jan: OK, we have the food producer, Robeki, who has no knowledge on Buddhist thoughts about to engage in a dialogue on Buddhist thoughts with a Buddhist monk. It looks like a sure victory. How possibly could there be any implications for game theory? By the way, are these two the only characters in the story? No beautiful princess?

Shinzuki: First of all, the jelly maker's name is Rokubei, not Robeki. Secondly, you should be patient. A lot of implications will be found after our story concludes. Finally, another character is involved. His name is Hachigoro and his role is to be a witness to the entire story. I'm sorry you are disappointed by finding no love stories or battles of the sexes included here.

Let me continue the story.

"Since Rokubei didn't know how to communicate with the monk on Buddhism, he didn't answer the monk's questions. The monk thought that being silent might be a style of dialogue, and tried asking questions in several different fashions. After some time, Rokubei began answering the monk's questions with hand gestures. Taking this as a style of dialogue, the Monk began responding with hand gestures as well. Eventually, both Rokubei and the monk agreed that Rokubei had defeated the monk. Then the monk left as a loser."

Jan: Why? I don't understand how Rokubei defeated the monk. What happened?

Shinzuki: I'm glad to hear you find the story strange. Perhaps here you will find the implications to game theory you are searching for. As you shall see, Hachigoro both witnesses the event and plays an important role by questioning the outcome. The concluding part of the story is as follows:

"After the dialogue, Hachigoro followed the monk and asked him about the dialogue. The monk answered that the master had expressed great Buddhist thoughts through his gestures and he should be respected. Hachigoro wondered about when, if ever, Rokubei had learned Buddhist thoughts, and he returned to the temple to ask him about it. Rokubei said he had never learned Buddhist thoughts. Rather, the monk started talking badly about his jelly products with his gestures. This angered Rokubei, and thus he beat the monk."

Jan: Wait a second. I understand the monk talked to Rokubei with gestures about Buddhist thoughts. Did the gestures by Rokubei have different meanings? What are they?

Shinzuki: In fact, in the full form of the story, Hachigoro heard the meanings of the gestures intended by the monk, as well as by Rokubei. For example, a gesture of creating a circle with one's arms means the universe to the monk, but it means a round Konnyaku product to Rokubei. All the gestures are meaningful to each person.

Jan: I understand the story itself, but where are the implications for game theory?

Shinzuki: All right, now, perhaps you are ready to understand the game theoretical implications. Each of Rokubei and the monk believed that they had a perfectly meaningful dialogue, and that the victory of Rokubei was common knowledge. However, the monk believed that they had a Buddhist dialogue, while Rokubei believed that they had discussed his jelly products.

Jan: Great. I understand your explanations about the story, but I'm still not able to see any implications for game theory. May I ask some questions?

[Without waiting for a response, Jan continues.]

First of all, did the gestures they exchanged become common knowledge? Secondly, is the victory of Rokubei common knowledge between them? I believe that your answer to these questions will be yes.

However, these people attached completely different meanings to their gestures. They don't share a basic understanding of what the gestures mean. In such a situation, can we say that they share the common knowledge of the victory of Rokubei as well as the gestures exchanged?

[Shinzuki becomes slightly overwhelmed.]

Shinzuki: How can you ask me such sharp questions? Let me think about them, uhum....

I would like to answer "Yes" to your first questions. The gestures exchanged and Rokubei's victory became common knowledge. Here, it is at least necessary to assume that their visions are normal. The speed of light is almost infinity, which is actually about 300,000 kilometers per second. Through their visions, they are able to verify in an instant, almost infinite observations of a gesture and the other's observation of that gesture. Hence, the common knowledge of the gestures and of Rokubei's victory, expressed and understood with gestures, is an adequate assumption. Incidentally, I found this argument while reading the "analogy of the sun" in Chapter 6 of Plato's *Republic* [23].

Now more to the point, the symbolic forms of the gestures can be assumed to be common knowledge. But these symbolic forms have completely different meanings to Rokubei and the monk. Therefore, the gestures and victory are common knowledge at a superficial level. In fact, they lead those people to profound misunderstandings of each other. Nevertheless, they still have the common knowledge of the gestures and victory.

Jan: Yes, you understand the point of my questions. When we talk about common knowledge or even about knowledge, the truthfulness is somehow assumed. In this story, the common knowledge of the gestures cannot be true since the two people attached different meanings to them.

Shinzuki: Isn't the common knowledge of the gestures true? Huhm..., I think your reasoning may be correct, but your conclusion is not true. The objects of common knowledge in our case are the symbolic forms of the gestures exchanged. The common knowledge here doesn't include the interpretations of the gestures. May I use another analogy?

Jan: Of course, but don't digress too much.

Shinzuki: A Japanese apple farmer would like to sell his apples, which are dear to him and should be eaten fresh. An American, visiting Japan, wants apples to bake apple pies. Without communicating their intentions, the exchange of apples and money occurs between these two people. With communication between the two, the exchange might not have occurred. But my point here is that apples are apples independent of the intentions attached to them by the farmer and the visitor. I would like to call this the "analogy of apples" in analogy with Plato's "analogy of the sun."

Jan: Once again, you have told a nice story, but where is the relevance to our discussion?

Shinzuki: I'm sorry, perhaps I should return to the truthfulness of the common knowledge of gestures. The truthfulness of the knowledge of a gesture is the accurate observation and recollection of the gesture. The intended interpretations of gestures are irrelevant just like the intentions of the buyers and sellers of apples.

Jan: Wow. You have brought out many interesting and puzzling aspects of the story. I will need some time to digest them. I should go home now and it is getting cold. Let's continue our discussion tomorrow morning. Then, could we please start talking about game theory?

Shinzuki: Yes, but tomorrow I will have a class until 10:30. Let's meet around 10:45. I will bring another heater.

2. False Beliefs and Decision Making

[In the lab, Jan has been waiting for Shinzuki, and then Shinzuki appears in a good mood carrying a heater.]

Jan: How was your class?

Shinzuki: Good, of course.

Jan: Last night I thought about the story more. Now I feel it may be related to game theory. I'm anxious about hearing the implications of the story for game theory. However, since you have a tendency to digress, I would like to impose a constraint on our discussions. We should directly discuss the implications to game theory and its relationship to the story. The use of analogies should be avoided.

Shinzuki: Huhm..., You are harsh, but I will try my best to play within your rules. Let's consider a very simple situation where players 1 and 2 play the prisoner's dilemma game. Could you please draw its payoff matrix on the blackboard and name it Table 2.0?

Jan: Sure, it is an easy task. I name the prisoner's dilemma game $g^0 = (g_1^0, g_2^0)$, since I expect you want to have other games.

Table 2.0: $g^0 = (g_1^0, g_2^0)$

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$
$\mathbf{s}_{11}$	$(5,5)$	$(1,6)$
$\mathbf{s}_{12}$	$(6,1)$	$(3,3)^*$

Shinzuki: Yes, you are right. I want to give two more games now, which I will write on the blackboard.

Table 2.1: $g^1 = (g_1^1, g_2^1)$

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$
$\mathbf{s}_{11}$	$(5,0)$	$(1,0)$
$\mathbf{s}_{12}$	$(6,0)$	$(3,3)^*$

Table 2.2: $g^2 = (g_1^2, g_2^2)$

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$
$\mathbf{s}_{11}$	$(0,5)$	$(0,6)$
$\mathbf{s}_{12}$	$(0,1)$	$(3,3)^*$

Now, suppose that player 1 thinks that the game is given as $g^1 = (g_1^1, g_2^1)$, while player 2 thinks that the game is $g^2 = (g_1^2, g_2^2)$. Each of them thinks about the common knowledge of a different game. Although the standard game theory literature assumes the common knowledge of the game structure, it would be apparently impossible to maintain this assumption in the current context. We modify this assumption to require that each player $i = 1, 2$ believes the common knowledge of the game $g^i = (g^i_1, g^i_2)$.

Jan: What? What do you mean by "Each player i believes the common knowledge of $g^i = (g^i_1, g^i_2)$"? Perhaps, I should verify if I understand you correctly. Huhm..., it is difficult to understand your statement. I should paraphrase it into a longer form: Player i personally believes that it is common knowledge that $g^i = (g^i_1, g^i_2)$ is the game to be played. Is this paraphrasing accurate?

Shinzuki: Yes, certainly.

Jan: All right. Now, it is getting clearer. The truthfulness of the common knowledge of $g^i = (g^i_1, g^i_2)$ is assumed only in the belief of player i. Is my understanding OK?

Shinzuki: You correctly understand my assumptions.

Jan: Great. Now, I should ask what kind of mathematical language you will use to represent your game theoretical argument. I would be very surprised if you talked about such a delicate problem without having any mathematical formulation.

Shinzuki: I will talk about that problem in the mathematical language called *epistemic logic*. I borrow the language of logic from [6] and [15].[2] Using this language, the game theoretical assumptions I uttered, and one logical conclusion derived from them are described as follows:

[Shinzuki writes on the blackboard.]:

(1) $\vdash g^0, \mathrm{B}_1\mathrm{C}(g^1), \mathrm{B}_2\mathrm{C}(g^2) \to \mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22}))$.[3]

I should give some explanations about these symbolic expressions. The first symbol, $\vdash$, means the entire statement following it is provable. The left-hand side following $\vdash$, and continuing before $\to$, consists of assumptions, or axioms. To the right of $\to$, we find another statement which is logically concluded from the left-hand side.

Jan: Let me interpret what you wrote on the blackboard. The first symbol in the left-hand side after $\vdash$ seems to mean that the game to be played is the prisoner's dilemma game g^0. The next two statements separated by commas seem to be that: player 1 believes[4] the common knowledge of g^1, and player 2 believes the common knowledge of

[2]Epistemic logic is a branch of modal logic. For a general introduction to modal logic, see Chellas [1] and Hughes-Cresswell [4]. Fagin *et al.* [2] and Meyer-van der Hoek [21] are introductory books dealing with epistemic logics.

[3]The provability $\vdash$ depends upon a logical system. In this paper, various logical systems are involved without specifying any. For details, see Kaneko [6] and Kaneko-Suzuki [15].

[4]In the game theory literature, belief is often understood as subjective probability in the sense of Savage [24]. The concept of belief here differs from subjective probability in many respects. The main difference is that we discuss beliefs in association with the logical abilities of players, while beliefs are treated as a black box given together with preferences in the Savage approach.

g^2. The right-hand side seems to mean that it is common knowledge that $(\mathbf{s}_{12}, \mathbf{s}_{22})$ is a Nash equilibrium. You said that the arrow means that the right-hand side is logically concluded from the left-hand side, didn't you? You also said that the symbol $\vdash$ means that the entire statement following it is provable. Isn't $\vdash$ redundant?

Shinzuki: Without the symbol $\vdash$, we only have the statement "if the left-hand side is assumed, then the right-hand side holds". The additional $\vdash$ means that the "if-then" clause is provable in an epistemic logic[5].

Jan: OK, OK. (1) states that it is provable that if g^0, $\mathrm{B}_1\mathrm{C}(g^1)$ and $\mathrm{B}_2\mathrm{C}(g^2)$, then the strategy pair $(\mathbf{s}_{12}, \mathbf{s}_{22})$ as a Nash equilibrium is common knowledge. It sounds alright, but the beliefs and knowledge are mixed. In the left-hand side, the beliefs of common knowledge are assumed, while in the right-hand side, the common knowledge of a Nash equilibrium is obtained. The beliefs seem to have somehow vanished and been replaced entirely by knowledge as if by magic. How did it happen?

Shinzuki: In fact, (1) is true. This is parallel to the Konnyaku Mondô in that both statements include false beliefs, but the conclusions are common knowledge. However, I admit that the uses of beliefs and knowledge are not parallel in (1). For a better understanding, I retreat here from (1) to the following slightly weaker assertion:

$$(2) \vdash g^0, \mathrm{B}_1\mathrm{C}(g^1), \mathrm{B}_2\mathrm{C}(g^2) \rightarrow \mathrm{B}_1\mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})) \wedge \mathrm{B}_2\mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})).$$

We shall discuss the relationship of (1) and (2) to the Konnyaku Mondô later.

Jan: Now, (2) is nicer than (1), since each "believes" the common knowledge of something in both the left-hand and right-hand sides. Huhm..., the belief of the common knowledge of a Nash equilibrium is derived from the individual belief of the common knowledge of the game g^i. This logical calculation takes place in the mind of each player. Then what is the role of the objective statement g^0?

Shinzuki: Oh my dear, Jan, I'm pleased to find you understand the problem perfectly. In fact, (2) is equivalent to the following two separated statements, neither of which involves g^0:

$$(2a) \vdash \mathrm{B}_1\mathrm{C}(g^1) \rightarrow \mathrm{B}_1\mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})).$$

$$(2b) \vdash \mathrm{B}_2\mathrm{C}(g^2) \rightarrow \mathrm{B}_2\mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})).$$

The derivation of (2a) and (2b) from (2) needs an advanced technique of logic, while the converse is just a calculation.

[5] Note that it may be the case that neither the "if-then" clause nor the "if-then not" are provable, i.e., they are undecidable. This may mean that a player can find neither some conclusion nor its negation from his basic beliefs (the left-hand side). The logic approach often evaluates such provability or unprovability.

Jan: It sounds wonderful, though I don't appreciate the difficulty in the derivation from (2) to (2a) and (2b). Now on the issue of game theoretical implications, consider the decision making of player 1. He believes that his payoff function is g_1^1, which is the same as his payoff function g_1^0 of the objective game g^0. One might argue that he chooses his strategy $\mathbf{s}_{12}$, because it is a dominant strategy rather than because it is part of a Nash equilibrium. Following this line of reasoning, we seem to be moving even further from common knowledge.

Shinzuki: All right, I can change (2a) into

$$\text{(3a)} \vdash \mathrm{B}_1(g^1) \rightarrow \mathrm{B}_1(\mathrm{Dom}_1(\mathbf{s}_{12})).$$

Jan: I'm surprised you changed your statement so easily. However, (3a) is in some sense nicer than (2a), since it includes no common knowledge. Huhm..., (3a) still includes player 2's payoff function in $\mathrm{B}_1(g^1)$. Can you change (3a) into the following?

$$\text{(3a}'\text{)} \vdash \mathrm{B}_1(g_1^1) \rightarrow \mathrm{B}_1(\mathrm{Dom}_1(\mathbf{s}_{12})).$$

Shinzuki: Yes, of course. Are you satisfied by having a purely personalized version?

Jan: Yes, I am very satisfied. Uhum..., it raises two opposite thoughts in my mind. On the one hand, (3a′) makes sense perfectly for me since player 1 simply derived $\mathbf{s}_{12}$ as a dominant strategy from the belief of his payoff function g_1^1. On the other hand, (3a′) no longer involves false beliefs since $g_1^1 = g_1^0$.

Shinzuki: Sorry. I change (3a′), again into

$$\text{(3a}''\text{)} \vdash \mathrm{B}_1(g_1^2) \rightarrow \mathrm{B}_1(\mathrm{Dom}_1(\mathbf{s}_{12})).$$

Please notice that I have changed the true payoff function $g_1^1 = g_1^0$ into the false payoff function g_1^2 of Table 2.2.

Jan: Once again you have flippantly adjusted the assumptions. I would like to understand the whole problem, instead of each small piece and it is difficult to concentrate when you keep changing assumptions. We seem to be moving further and further away from our initial objective. While I applaud you for refraining from the use of analogies, I must admit you are weaving an impossible web in my head with your incessant manipulation of assumptions. Please, stop and return to something concrete. I almost forget the point.

Shinzuki: Aha, I return to a full statement now:

$$\text{(3)} \vdash g^0,\mathrm{B}_1(g_1^2),\mathrm{B}_2\mathrm{C}(g_1^2, g_2^1) \rightarrow \mathrm{B}_1(\mathrm{Dom}_1(\mathbf{s}_{12}))\wedge\mathrm{B}_2\mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})).$$

Surely, we can talk about the falsity of beliefs of players in (3).

Jan: Although this looks slightly different from our starting point, it does seem to be moving back in the right direction. Let me try to make sense of (3). I start with the left-hand side of $\rightarrow$. Uhum..., this involves a lot of falsities. Probably, for my own sanity, I should translate (3) into English. First, it says the objective situation is described by the prisoner's dilemma game g^0. Next, it says that player 1 believes his payoff function is given by g_1^2, which we know to be false relative to g^0. Next, it says that player 2 believes the game (g_1^2, g_2^1) is common knowledge between the players, which we know to be false both relative to the objective game g^0 and the content of player 1's belief.

In this formulation, player 1 ignores, or does not discuss, player 2's payoff in the formation of his beliefs, while player 2 believes they have common knowledge of (g_1^2, g_2^1) which we know to be false relative to g^0. Is my understanding of the left-hand side of (3) accurate?

Shinzuki: Yes, it is.

Jan: Good. Now let's attack the right-hand side. The first piece is the derivation by player 1 that his strategy $\mathbf{s}_{12}$ is a dominant strategy. This derived statement happens to be objectively true for the game g^0. We also, find the derivation by player 2 that the pair $(\mathbf{s}_{12}, \mathbf{s}_{22})$ as a Nash equilibrium is common knowledge. The pair $(\mathbf{s}_{12}, \mathbf{s}_{22})$ to be a Nash equilibrium is also true relative to g^0. However, player 2 believes that this fact is common knowledge between 1 and 2, while player 1 does not think about player 2 at all. This is yet another kind of falsity.

Shinzuki: Indeed, we meet a lot of falsities in beliefs, in particular, both sides of (3) involve falsities. Actually, I deliberately wrote (3) in a puzzling form.

[Jan answers in a tone of disagreement.]

Jan: Yes, (3) has a lot of falsities. Huhm..., however, the falsities on both sides of the arrow $\rightarrow$ are parallel. The previous asymmetry of falsities in (1) is, in some sense, more interesting, since the true common knowledge is derived from the false beliefs of common knowledge. As a game theorist, I would typically assume that the objective truths on the right-hand side were obtained from objective truths in the minds of the players. You have shown that this need not be the case. Is this distinction important for game theory? Can you connect this finding in (1) to the Konnyaku Mondô discussed yesterday?

But now my head is reeling..., and noises are rising within me.

Shinzuki: I think you are saying you are hungry. Why don't we go to the Mexican Restaurant? The lunches there are cheap and large. Although I would like to continue our discussions right after lunch, I will have another stupid meeting until 3PM. Is it convenient for you to have more discussions after that?

Jan: Yes, of course. Anyway, it is time to eat. Let's go.

3. More on Decision Making

[After lunch, Shinzuki went to a meeting, and Jan has been taking a nap on a couch in the lab.]

Shinzuki: Hi Jan. I'm back from my "meeting from hell." Ah, I'm sorry to wake you up.

Jan: No, no, thank you for waking me up. I wanted to wake up by myself. I was very sleepy because of the intense discussions this morning and the good food in the restaurant. How was your meeting?

Shinzuki: Nothing important as usual. Actually, I didn't sleep well in the meeting because the chairperson spoke so loudly.

Jan: Ha ha, you are in almost the same state as me. By the way, I thought about our discussions during my nap. Your explanations diverged quite a lot.

Shinzuki: I have a more general statement including the previous ones, which I will write on the blackboard as:

$$(4) \vdash \Gamma^0, \mathrm{B}_1(\Gamma_1), \mathrm{B}_2(\Gamma_2) \rightarrow \mathrm{B}_1(\mathrm{D}_1(s_1)) \wedge \mathrm{B}_2(\mathrm{D}_2(s_2)).$$

Here, the first assumption set Γ^0 is the objective description of the situation, and $\mathrm{B}_i(\Gamma_i)$ is the beliefs owned by player $i = 1, 2$. The right-hand side means that each player i derives his decision s_i. Again, the symbol $\vdash$ means that the entire sentence is provable.

Jan: Shinzuki, please stop for a moment and let me digest your current statement. First, I like the fact that it does not seem to involve common knowledge at all. Our previous discussions led me to find that common knowledge is only in the mind of a player. This suggests that we can do away with common knowledge operators and focus only on beliefs. This appears to be what you have done. However, you also replaced g^0 by Γ^0, and that worries me. If you included some "common knowledge" there, then I would not be so happy.

Shinzuki: We can suppose that Γ^0 contains purely objective statements and no instances of "common knowledge" or beliefs.

Jan: Good. Then I can go further and try to understand (4) now. Let me put (1), (2), (3), and (4) next to each other on the blackboard.

$$(1) \vdash g^0, \mathrm{B}_1\mathrm{C}(g^1), \mathrm{B}_2\mathrm{C}(g^2) \rightarrow \mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})).$$

$$(2) \vdash g^0, \mathrm{B}_1\mathrm{C}(g^1), \mathrm{B}_2\mathrm{C}(g^2) \rightarrow \mathrm{B}_1\mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})) \wedge \mathrm{B}_2\mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})).$$

$$(3) \vdash g^0, \mathrm{B}_1(g_1^2), \mathrm{B}_2\mathrm{C}(g_1^2, g_2^1) \rightarrow \mathrm{B}_1(\mathrm{Dom}_1(\mathbf{s}_{12})) \wedge \mathrm{B}_2\mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})).$$

(4) $\vdash \Gamma^0, \mathrm{B}_1(\Gamma_1), \mathrm{B}_2(\Gamma_2) \to \mathrm{B}_1(\mathrm{D}_1(s_1)) \wedge \mathrm{B}_2(\mathrm{D}_2(s_2))$.

Mathematically speaking, both (2) and (3) appear to be special cases of (4), but (1) is not.

Shinzuki: Uhum.., (1) is not a special case. However, we can include (1) by changing (4) slightly.

Jan: Ha, ha, ha. I don't care about generality. Rather, I'm curious about the treatment of strategies here. I find that in (4) each player i has a decision s_i, while in (1) and (2) each player thinks about the Nash strategy pair $(\mathbf{s}_{12}, \mathbf{s}_{22})$, and in (3), player 1 thinks about $\mathbf{s}_{12}$, but 2 thinks about $(\mathbf{s}_{12}, \mathbf{s}_{22})$. In this sense, they are asymmetric. More specifically, in (3) $\mathbf{s}_{12}$ is 1's decision, but what is $(\mathbf{s}_{12}, \mathbf{s}_{22})$ for player 2? I ask this because player 2 cannot choose $\mathbf{s}_{12}$.

Shinzuki: That is a good question, since I know the answer. The strategy $\mathbf{s}_{12}$ appears as a prediction of 1's decision making in 2's mind. Player 2 doesn't choose $\mathbf{s}_{12}$, but he thinks about 1's decision making since he requires it to predict 1's decision making. To a large extent, this argument is rather standard from the game theoretical analysis of Nash [22].

Jan: But if (4) is a generalization of (2) and (3), then why doesn't it include predictions?

Shinzuki: Aha, (4) may include predictions in $\mathrm{D}_i(s_i)$, though they are not explicitly written. In the case of (3a), since a dominant strategy is the decision criterion, the prediction part is not included at all.

Jan: OK, so again we move further from common knowledge. Do we really need common knowledge?

[Shinzuki speaks now in an authoritative tone.]

Shinzuki: Your question is naive. The issue has many aspects. First, we need to think about the treatment of common knowledge and common beliefs, and more fundamentally, the difference between knowledge and belief. Second, we should think about the relationship between common knowledge and decision making. We found by our analysis of (3) where player 1 has a dominant strategy, that common knowledge is "not necessarily" needed for decision making. Are you interested to find when common knowledge is necessary for decision making?

Jan: I would like to hear about the distinction between knowledge and belief now, but OK, since you seem to want to, why don't you tell me about when common knowledge is needed?

[Shinzuki changes his tone, speaking with a German accent.]

Shinzuki: Yah, yah..., I promise to discuss the distinction between knowledge and belief later.

[In normal voice]

But for now, let's discuss when common knowledge is needed. You will find that the argument is amazingly standard in the game theory literature. Consider the following form of prediction-decision making:

(a) 1 maximizes his payoff predicting that 2's decision is governed by (b);

(b) 2 maximizes his payoff predicting that 1's decision is governed by (a).

Jan: Hold on..., (a) appears in (b), and (b) appears in (a). This seems to involve circular reasoning. The sentence (a) can be plugged into (b), and (b) into (a), and so on, forever. This process yields an infinite regress, and I have a great head ache with such circularities and infinite regress. Is there any meaningful solution?

Shinzuki: Your concern is valid. To treat that infinite regress, we need a common knowledge extension of an epistemic logic. In such an extension, we find a complete solution which turns out to be the common knowledge of Nash equilibrium. If (a) and (b) are restricted to occur in the mind of a single player, then the solution is the personal belief of the common knowledge of Nash equilibrium, i.e., $\mathrm{B}_i\mathrm{C}(\mathrm{Nash}(s_1, s_2))$.[6]

[In an authoritative tone]

Exactly speaking, (a) and (b) are properties required for a type of decision criterion. It is my claim that if such a decision criterion is adopted, then common knowledge is necessarily involved.

[Shinzuki starts speaking in a low voice with a critical tone.]

Some people argue that common knowledge is unnecessary for Nash equilibrium. But that is not the issue! The issue is the necessity of common knowledge for decision making, not Nash equilibrium. Nash equilibrium is Nash equilibrium, as an apple is an apple in the analogy of apples. Nash equilibrium should be distinguished from the description of decision making.

Jan: You seem to mean that the necessity of common knowledge may depend upon the decision criterion used. Hummm,....the dominant strategy criterion does not involve common knowledge. However, it might not enable a player to make a decision in a game. Is there any decision criterion that always gives a decision and does not involve common knowledge?

Shinzuki: Actually, there is the simplest example called the *default decision criterion.* It recommends a pre-specified strategy, say, the first strategy, without requiring the decision maker to think about anything. If you think it is simple, I can give more

[6] See Kaneko [6].

examples based on this.

Jan: I'm getting tired..., but the last point sounds very nice. Maybe we should have a coffee break. But before that, I have one thought I would like to articulate. Common knowledge involves an infinite regress, or infinite depths of nested beliefs structures. On the other hand, if the dominant strategy decision criterion or the default criterion is adopted, then no common knowledge is required, since there is no nesting of beliefs, or the depth of beliefs is only 1.

Shinzuki: Yes, you are correct. Would you like to see an example involving a finite depth of beliefs greater than 1 without using the default criterion?

Jan: Well, I'm tempted by coffee, but please go on.

Shinzuki: The example is very simple and obtained from changing only one payoff of player 2 in the prisoner's dilemma:

Table 2.0$'$: $g^{0'} = (g_1^{0'}, g_2^{0'})$

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$
$\mathbf{s}_{11}$	$(5,5)$	$(1,\underline{2})$
$\mathbf{s}_{12}$	$(6,1)$	$(3,3)^*$

In this game, player 2 has no dominant strategy, but if player 2 believes that player 1 follows the dominant strategy criterion, then 2 predicts $\mathbf{s}_{12}$ as 1's decision, and based on this prediction, he chooses $\mathbf{s}_{22}$. This is written as

$$(5\text{b}) \vdash \mathrm{B}_2(\mathrm{g}_2^{0'} \cup \mathrm{B}_1(g_1^{0'})) \to \mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(\mathbf{s}_{12})) \wedge \mathrm{Best}_2(\mathbf{s}_{22} \mid \mathbf{s}_{12})).$$

I didn't write 1's statement because it is too long.

Jan: OK, $\mathrm{B}_1(\mathrm{Dom}_1(\mathbf{s}_{12}))$ in the scope of B_2 is the prediction of player 2 about 1's decision making. The depth of nested beliefs is 2 here, isn't it? Then, $\mathrm{Best}_2(\mathbf{s}_{22} \mid \mathbf{s}_{12})$ means $\mathbf{s}_{22}$ is the best strategy to the prediction $\mathbf{s}_{12}$. It all makes sense. I think, I can now find more examples which require finite depths of nested belief structures. This appears to capture some aspect of bounded rationality, doesn't it?

Shinzuki: Yes, it is one aspect of bounded rationality. You will find more examples in [15].

Jan: Now I remember how tired I am. I want to ask one more small question before coffee, since I feel we have reached some deep implications for game theory. My last question is about the multiplicity of... uhum..., what?

Shinzuki: You are really tired. I think you want to say the multiplicity of candidates for decisions.

Jan: That's right, in fact your wording is better than what I was thinking. Suppose we have a game like the battle of the sexes with two Nash equilibria:

Table 3.1: $g^4 = (g_1^4, g_2^4)$

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$
$\mathbf{s}_{11}$	$(2,1)^*$	$(0,0)$
$\mathbf{s}_{12}$	$(0,0)$	$(1,2)^*$

Even if each player has a belief of common knowledge of a Nash equilibrium, each may have a different Nash equilibrium in mind. Then it might seem that each can make a decision to use his equilibrium strategy. But the combination is not a Nash equilibrium. How do you treat such a situation?

Shinzuki: You need an additional assumption such as the common knowledge of one Nash equilibrium, for example, $(\mathbf{s}_{11}, \mathbf{s}_{21})$. This may be included in Γ_i of (4). If both players have the same common knowledge assumption, then the resulting outcome is $(\mathbf{s}_{11}, \mathbf{s}_{21})$.[7] However, if only player 1 has this, then the outcome $(\mathbf{s}_{11}, \mathbf{s}_{21})$ may not be reached, yet player 1 chooses $\mathbf{s}_{11}$ and predicts $\mathbf{s}_{12}$ to be chosen by player 2. To obtain such common beliefs in one strategy combination needs communication, and such communication is external to epistemic logic.

Jan: Let's go for coffee now, to avoid more divergence. Do you know any good places?

Shinzuki: That is a very difficult question. Anyway, we can get coffee in the village.

4. General Principles

[Shinzuki and Jan are walking back from a coffee shop.]

Jan: It was good coffee, but the coffee shop was too crowded. I have an interesting idea about how to control the number of customers. Can you imagine it?

Shinzuki: Yes, I can, of course. It is that the price should be increased according to the economic principle that when price rises, demand will fall.

Jan: That is one way. Another way is to decrease the quality of the coffee. Then some customers would leave.

Shinzuki: Ha ha, who makes such a stupid decision?

Jan: Yes, yes, a coffee shop in my town did this really. It succeeded in decreasing the number of customers. I myself went there only once after the reduction in quality. Can you believe this control?

Shinzuki: It is difficult to believe. Now I should impose a constraint on you not to digress from our main problem.

[7] See Nash [22] and Luce-Raiffa [18] for interchangeability of Nash equilibria, and also Kaneko [5] for the further treatment of games with/without interchangeablity from the viewpoint of individual decision making.

Jan: Sorry. In fact, this is my revenge for your past digressions!

Shinzuki: Indeed, it is a nice lesson in understanding the transgression of digression. I will think about how to get you back.

Jan: Now, let me try to recall what should be discussed. Uhum..., you promised to discuss the distinction between knowledge and belief.

Shinzuki: Good, I remember our problem, too. In an European tradition of philosophy it is standard to define

(6): knowledge is a justified true belief[8].

In this definition, belief is more basic than knowledge. For the moment, please take a "belief" as something held in the mind of the believer. I want to talk about the word "true" in (6). This notion of truth is taken from the perspective of the objective observer, or at least some thinker other than the one holding the belief.

Jan: OK, you are saying that the truth in (6) is external to the believer. There might also be a truth included within the belief of a person. But this truth is beyond dispute for him and we cannot discuss the truth or falsity of such beliefs within his perspective. We can only objectively distinguish between false and true beliefs if we consider them from an external perspective. A true belief is a candidate for knowledge, isn't it? I think this is compatible with what we discussed before.

[Jan and Shinzuki arrive at the laboratory.]

Shinzuki: Very good. In my previous statements, such as (2) and (3), there are at least three references to truth. One is from the perspective of the outside objective observer, and then there is one from the perspective of each player. In (2), g^0 is the situation true for the outside observer. From this perspective, the content, g^i, of $\mathrm{B}_i\mathrm{C}(g^i)$ is false, but, from the perspective of each player i his belief $\mathrm{B}_i\mathrm{C}(g^i)$ is true and beyond dispute. We discussed the falsities involved in (3) already.

Jan: It might help my understanding to consider the case where all beliefs are true even from the perspective of the outside observer. Consider the following:

(7): $\vdash g^0, \mathrm{B}_1\mathrm{C}(g^0), \mathrm{B}_2\mathrm{C}(g^0) \rightarrow ??$

I put ?? in the right-hand side since the conclusion part is irrelevant to the point I want to make now. According to your explanation, the assumption part of (7) is indistinguishable from the common knowledge of g^0. Is there any meaningful difference

[8] The definition of knowledge has been debated for a long time (see papers in Moser [20]). We interpret the terms "justified" and "truth" as requirements to be checked from the perspective of an external investigator.

between this and simply writing $C(g^0)$?

Shinzuki: Actually, it can be proved that the set g^0,$B_1C(g^0)$,$B_2C(g^0)$ is equivalent to $C(g^0)$, where the notation g^0,$B_1C(g^0)$,$B_2C(g^0)$ is an abbreviation of $g^0 \cup B_1C(g^0) \cup B_2C(g^0)$. Moreover, this is equivalent to $CC(g^0)$, and to $CC...C(g^0)$ with any finite number of applications of C.[9]

Jan: Really? That is interesting. In this case, everything is OK with me and it also seems compatible with the standard "implicit" game theoretical arguments. With this new understanding, I would like to consider more cases of false beliefs. Let's return to the assumption part, g^0,$B_1C(g^1)$,$B_2C(g^2)$, of (2). Because there are false beliefs involved, it seems possible that one of the players might notice this inconsistency and the assumption set would fall apart. How do you guarantee that the assumption set is compatible and might persist?

Shinzuki: I'm glad you are now getting through the gate of epistemic logic.

Jan: You are flattering me, but I'm still pleased with my finding, and perhaps too at your noticing my finding. Anyway, please continue.

Shinzuki: OK. The assumption set g^0, g^1, g^2 is inconsistent, and so is the set g^0,$C(g^1)$, $C(g^2)$.[10] If we start with either of these assumption sets, the theory is nonsense since anything, including absurdities, can be derived from a set of inconsistent assumptions. To prevent such absurdities, and allow the persistence of a set of assumptions, it suffices to prove that the set is consistent. If consistency is shown, then no player will ever notice an inconsistency and all will be fine. This is really the starting point of the research in the fields of "epistemic logics and game theory". We use some developments from proof theory and model theory to guarantee the consistency of g^0,$B_1C(g^1)$,$B_2C(g^2)$ which can be found in [6].

One notable finding is that to obtain the consistency of this set, we need to drop one famous axiom from our epistemic logic:

(8): $B_i(A) \rightarrow A$.

This is often called the *truth axiom* or *veridicality axiom*. If (8) is assumed in an epistemic logic, then every belief is assumed to be true in the eyes of the outside observer, and false beliefs cannot be discussed.

Jan: Sorry to interrupt you, but another point which I want to mention now is about

[9] There are two approaches to extending epistemic logic with common knowledge: the fixed-point approach and the infinitary approach. For the former, see Fagin *et al.* [2] and Meyer-van der Hoek [21]. For the latter as well as the relationship to the former, see Kaneko *et al.* [14].

[10] Throughout this paper, we assume that $C(A)$ means the common knowledge of A rather than the common belief of A. In our context, the common belief of A is definable as $B_1C(A) \wedge B_2C(A)$. Conversely, if we start with the common belief $C_B(A)$ as a primitive, then the common knowledge of A is definable as $C_B(A) \wedge A$. See Kaneko *et al.* [14] as well as Section 5 of Kaneko [6].

the meaning or interpretation of the consistency of an assumption set, and the viewpoint from which it is derived.

I can now think of four meaningful different perspectives and assumption sets. The assumption set for player 1 is clearly $\mathrm{B}_1\mathrm{C}(g^1)$, and for player 2 it is $\mathrm{B}_2\mathrm{C}(g^2)$. The last two perspectives involve the outside observer who may have either g^0 alone or the entire set g^0,$\mathrm{B}_1\mathrm{C}(g^1)$,$\mathrm{B}_2\mathrm{C}(g^2)$. While we can imagine an outside observer who knows only the true game g^0, we can also imagine an objective observer who knows all of g^0,$\mathrm{B}_1\mathrm{C}(g^1)$,$\mathrm{B}_2\mathrm{C}(g^2)$. This second type of objective observer is rather close to the position I am taking in my analysis, and the one you seem to be taking in your analysis.

Shinzuki: Your last distinction is very nice. The consistency of g^0,$\mathrm{B}_1\mathrm{C}(g^1)$,$\mathrm{B}_2\mathrm{C}(g^2)$ is proved from the viewpoint of the second objective observer you mentioned, even though I may have seemed to take the perspective of the first observer.

[Jan is pleased.]

Jan: I'm slightly curious about how to prove the consistency of g^0,$\mathrm{B}_1\mathrm{C}(g^1)$, $\mathrm{B}_2\mathrm{C}(g^2)$, and perhaps this will help me understand what I just tried to describe. Can you give a brief explanation of the method of proof?

Shinzuki: Yes, I can. A proof needs a lot of details. However, believing the consistency is much like believing the logical possibility of the Konnyaku Mondô. The Konnyaku Mondô involves a set of false beliefs that is just like the assumption set g^0,$\mathrm{B}_1\mathrm{C}(g^1)$,$\mathrm{B}_2\mathrm{C}(g^2)$. By admitting the logical possibility of the Konnyaku Mondô, you also admit a belief in the possibility of its consistency. When you heard the story, you most likely searched for some way to make sense of the story. This is what a proof of consistency does. The situation arrived at for making sense of the story, is called a *model*. If we can construct a model of g^0,$\mathrm{B}_1\mathrm{C}(g^1)$,$\mathrm{B}_2\mathrm{C}(g^2)$, then we can use the *soundness theorem* to obtain the consistency of g^0,$\mathrm{B}_1\mathrm{C}(g^1)$,$\mathrm{B}_2\mathrm{C}(g^2)$.

Jan: Stop, stop. It seems a lot of things are needed for that step. Please postpone your explanation of the construction of a model and the soundness theorem to sometime in the future. I am learning a lot now from your description and by thinking about the story in this new light. Can we change our direction slightly and discuss the word "justified" in (6).

[Shinzuki shows his disappointment by raising his hands to his head.]

Shinzuki: Arghhh..., I'm slightly disappointed, since that construction is my favorite part. Perhaps later you will appreciate it more. OK, let's consider "justified" in (6). It is not very exciting, technically speaking. Usually, "justified" is quite ambiguous, but it is clear-cut in the epistemic logic approach since it unearths everything. By a justified belief, say A, for player i, we mean that player i has some argument, or justification for A from his basic beliefs. The basic beliefs are written in Γ_i of (4). For example, if player i has a mathematical proof for A from Γ_i, then this A is a justified belief.

Jan: It is clear-cut. A belief is justified when, for example, a player has a proof of it, uhum..., this proof is taken from his basic beliefs, right? So shouldn't we ask for a justification for a basic belief?

Shinzuki: No, basic beliefs are not justified in this sense.

Jan: But you claimed that the epistemic logic unearths everything.

Shinzuki: Sorry. I should say it unearths everything that can be unearthed.

[A bit surprised by Shinzuki's statement, Jan retorts.]

Jan: Ha ha, that sounds like a tautology. Is it everything or nothing? It seems to me that epistemic logic refuses to dig.

Shinzuki: You are right. I retreat by saying epistemic logic unearths everything before the consideration of basic beliefs Γ_i. We need to look for some other sources for basic beliefs.

Jan: What other sources for basic beliefs do you have in mind?

Shinzuki: This part is not well developed, though it has been discussed since the age of Plato. One source is one's experiences. This process is often called "induction" and it means to obtain a general law from finite experiences. Please note that this "induction" differs from the mathematical induction principle. While the latter principle has the same name, it is essentially a principle of deduction. In fact, the latter may be included in Γ_i,[11] but the former is external to the logical framework and is the process of constructing and revising the basic beliefs in Γ_i.[12]

Jan: I'm surprised to hear that the mathematical induction principle is one of deduction. The other "induction" you speak of is not deductive at all. When I learned the distinction between induction and deduction at university, both were described as processes of scientific reasoning.

Shinzuki: One difference here, from what you learned, is that our target is a human inductive process rather than a scientist's inductive one. The latter requires careful statistical treatments, but the former is rather a bold process since a great generalization is often made from a few experiences. We all have this tendency to generalize. For example, you went to the coffee shop only once after it decreased the coffee quality, but the shop might have increased the quality later. Are you sure that the coffee quality there is still bad?

Jan: Touche! I understand your distinction between human and scientific inductive processes.

[11] To include the mathematical induction principle in Γ_i, we need predicate epistemic logics. For such predicate epistemic logics including common knowledge, see Kaneko *et al.* [14].

[12] The revision of beliefs has been discussed in the literature on belief revision. See Schulte [25] for a good survey.

Shinzuki: I would like to give one more example in this line of reasoning to deepen your understanding. Suppose that you meet two people of an ethnic group that is new to you. You observe differences in their appearances as well as their behavior relative to the standard in your community. You might conclude that all people of that ethnicity have similar appearances and behavior.

Jan: I understand that these generalizations are rampant in our lives, but on the positive side, they allow us to expand and revise our basic belief set Γ_i.

Shinzuki: Your point is taken, but I would rather like to point out that once again we have arrived at some false beliefs. The generalities obtained by induction involve false beliefs in the sense that they are not valid from an objective viewpoint. These false beliefs are related to sociological problems of discrimination and prejudice. The inductive method I have been speaking of is formulated in [13], which the authors call the *inductive game theory*. In that paper, the emergence of racial prejudices is discussed, but connections to epistemic logic are not yet fully reached.

Jan: It sounds interesting. I will look at the paper. However, I doubt I can attribute all my beliefs to my past experiences. My experiences are very limited, yet I seem to have a limitless fountain of beliefs. I even have some on things I never experienced, like what it would be like to fly.

Shinzuki: Yes, experiences are limited for an individual being, but we also have communication and education. The experiences of former generations have been communicated to later generations through discussions, books, and other sources. A lot of "trial and error" has been taken by our ancestors, and many of these experiences have been passed on. For example, our distinction between poisonous and edible mushrooms is possible only after many trials and some fatal errors by previous generations.

Jan: The accumulation of such experiences are taught to us. Great! We have been making progress by our experiences.

Shinzuki: You are right to the extent that what we have discussed is quite a standard view of progression. However, I want to point out that communication and education are powerful, but have their own limitations. For communication and learning as well as for one's own memory of experiences, language is a useful and a necessary tool. Language consists of primitive symbols and grammars, just as a logical system consists of symbolic expressions and inference rules. In this sense, a logical system can be regarded as the idealization of a language.

[Having tuned out slightly, Jan notices Shinzuki is staring at him waiting for a reply.]

Jan: What is your point?

Shinzuki: My point is as follows. Suppose somebody has experiences. However, he translates his experiences into language in order to communicate those experiences to others or even just to memorize them for himself. In this translation, something is lost as well as gained. Something is lost by the fact that the communicator must choose some part of his experience in order to be understood by others. Something is gained by adding the structure required to express the experience in a language understandable by others. Raw experiences can never be transferred to other people.

Jan: It would be nice if you could relate it directly to what we have been discussing.

Shinzuki: OK, I will continue. People having experiences construct a simple story explaining experiences, and other people are taught this story through communication or education. This is a source of the basic beliefs Γ_i of (4).

[Shinzuki's voice becomes louder.]

Now, I stress the story is coherent, but it must also be simple so it can be memorized or communicated. This implies that Γ_i is innately subject to some falsities. If the story is extremely accurate and has every element of experience, it is useless since it is overwhelmingly long. Also, everybody has a conscious or unconscious tendency to adjust the story so it is more comfortable for him or her. Such adjustments were described in a funny, but sad, manner in Akira Kurosawa's movie "Rashoumon".[13]

[Shinzuki is roaring with enthusiasm, and Jan feels slightly taken aback.]

Jan: I understand what you want to say. My view of progress may be too naive. Does your argument have any relationship to the Konnyaku Mondô?

Shinzuki: Yes, it does! You are finally becoming very sensitive to the contradictory elements in our discussions. In the Konnyaku Mondô, the exchange of gestures is a communication. It represents the symbolic nature of language. Each person constructs a story from the gestures, which includes falsities. This example is an extreme one, but it shows that communication is never entirely free from fallacious elements.

Jan: Yes, indeed, but still a lot can be communicated by discussions like the ones we have been having.

Shinzuki: I'm afraid we might be having nothing more than a Konnyaku Mondô. Just like you, I believe we have had meaningful discussions, but in actuality, the whole time I might have been talking about Konnyaku products in logical terms..., and you, my friend, might have been discussing the battle of the sexes in game theoretical terms.

Jan: You are too cynical! I think we should stop for today. Why don't we go to the public bath tonight? The public bath increases the quality of our life here very much. I hope the town never decides to decrease the quality of the water to control crowding problems!

Shinzuki: Ha, ha, it is a great idea to meet at the public bath. It helps you to understand the Japanese culture, doesn't it? Shall we meet there around 8pm?

[13] The original short story for the movie "Rashoumon" is called "Yabunonaka" by Ryunosuke Akutagawa. He wrote a different story with the title "Rashoumon". See Kaneko [7].

Dialogue II: 8 Years Later

[Setting: July 2010, Kurai Shinzuki is visiting Jan Hammer at a university in Queensland, Australia close to the beach. Since Jan's visit in 2002, they have worked hard and developed inductive game theory. This morning, after a long walk on the beach, they return to Jan's office. Part II takes place in Jan's office, which is cluttered by papers and books.]

5. Information Pieces and Memory Functions

Jan: It is hard to imagine we started working on inductive game theory (IGT) 8 years ago. Since then, we worked hard, and very fortunately, we achieved a lot. Did you see the file I sent last night? As I mentioned to you on the beach, my new secretary found the recorded file of our early discussions. She seemed to be amused by Konnyaku Mondô.

I was surprised to find how much of our development in IGT was foretold in those early discussions.

Shinzuki: Yes, I saw the file, and indeed, I was surprised too. I'm also puzzled by it, since it predicted our later developments well. After those discussions, we spent so much time making trials and failures. I regret that we didn't return earlier to the discussions for some guidance.

Jan: I regret it, too.

Shinzuki: We say "What's done cannot be undone." We should forget it, no no, sorry. We should forget our regret, but we shouldn't forget our thoughts, trials, and failures.

Jan: True. Perhaps, we should try and recall what we did, and why it took so long to get here.

Do you remember how we planned to consider IGT using extensive games of von Neumann-Morgenstern [29] and Kuhn [17]? Then, we got stuck.

Shinzuki: Yes, I remember it well. Let's summarize the new concepts we introduced. Then, we may recall the reasons for our trials and failures.

Jan: Do you need to use such a negative expression? It sounds like you aimed at failure.

Shinzuki: Perhaps I should say that I made a lot of trials not aiming at failure, but objectively I succeeded in achieving failures, ha ha.

Jan: I think it is better to discuss positive problems.

Let us list the concepts we developed in the paper [9] called the "basic scenario". Do you remember how we needed to change even the most basic components in an extensive game? It started with our attempt to dump the full cognizance assumption.

[Jan writes on the whiteboard.]

Full-Cognizance Assumption (FC): each player knows the entire extensive game structure including information sets and payoff functions.

Many people write that the game is common knowledge, but they don't seem to appreciate that the heart of the problem is not common knowledge but full cognizance.

Shinzuki: You are right. Our target is how a player acquires structural knowledge on his social situation. For this, we take the opposite starting point to the FC assumption:

[Shinzuki now writes on the whiteboard.]

No-Knowledge Assumption (NK): each player has no *a priori* knowledge about the extensive game structure.

By turning FC on its head to obtain NK, the question of how a player acquires structural knowledge from his experiences starts to make sense. On the other hand, under FC, this question doesn't make sense at all, since FC presumes full structural knowledge already.

Jan: It is my turn to talk. Here is a critical point here. The standard formulation of an extensive game rests upon FC. More specifically, the concept of an information set doesn't make sense without it.

Shinzuki: It is a critical point. Can you give a more explanation of what you are talking about?

Jan: Of course, I can, since I'm older and more mature than I was 8 years ago. Ah, like Bob Dylan, I should say "I'm younger than that now." Indeed, I feel mentally younger than 8 years ago, and physically, um...

Shinzuki: Please stop quoting Dylan, and explain the problem of an information set.

Jan: Okay, the concept of an information set is used to describe a situation in which a player receives the same information at two or more nodes in the game tree. When the information u is written as a set of nodes like $\{x_2, x_3\}$ in Fig.5.1 I'm drawing on the whiteboard, PL1 must be cognizant of the underlying game structure. Otherwise, he cannot understand the expression $\{x_2, x_3\}$.

Shinzuki: Very good. The interpretation and description of an extensive game seems to work well with FC. Then, does this game situation disappear if we replace FC by NK? It still exists, doesn't it? What is wrong?

[Jan is a bit troubled.]

Jan: Hmmm. You can explain such philosophical problems better than me. I'm biologically more fit for asking questions than for providing answers, perhaps because my ancestors were naturally selected for their asking, not answering, ha ha ha.

Shinzuki: Okay, I will take control.

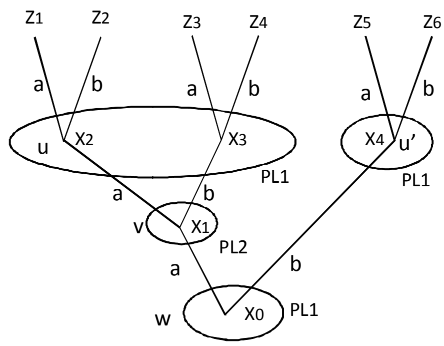

Figure 5.1: An Extensive Game

First, I claim that the concept of an information set should be defined from the objective point of view and should describe the fact that a player receives the same "information" at any node in one information set. In Fig.5.1, PL1 receives the same "information" at x_2 and x_3. This fact remains under either NK or FC, since receiving "information" is an objective phenomenon. Hence, the situation described by Fig.5.1 still remains even if we replace FC by NK.

Under FC, receiving "information" u implies that PL1 can infer that either x_2 or x_3 is now occurring but he does not know which. Hence, $u = \{x_2, x_3\}$ holds for him. On the other hand, under NK, we can still say "PL1 receives information u", but he cannot decompose u into $\{x_2, x_3\}$, since he doesn't know the game structure at all. Thus, under NK, we shouldn't call u an information set. That is why we replaced an "information set" by an "*information piece*" in [9].

Jan: We should question what the "information" u is exactly. An "information piece" is literally a piece of information; for example, it is simply a symbol like a color, or an expression in our natural language, or a well-formed formula in the sense of logic. For standard game theorists, the information piece or set is, more or less, the same. For us, the information piece has no set-theoretical contents in the beginning; I can receive the information piece "it is raining" without conceiving of any corresponding information set.

Shinzuki: Good! By treating "information" as pieces received, we can handle both FC and NK within extensive games. Under FC, each player can derive the set-theoretical content of an information piece u by constructing the set of nodes where u is received. FC is handled in this manner. Under NK, the information pieces remain symbolic without any set theoretical extension or possibly any meaning for a player.

Jan: This simple change from information sets to information pieces allows us to handle both FC and NK. However, the concept of information sets playes another role in the standard extensive game theory. The memory of a player is also described by an information set. For example, in Fig.5.1, the information sets u and u' are distinguished by PL1, since he receives different information at u and u' since each includes his memory of having taken a different action previously at w.

This is too bad; the standard approach with information sets mixes memory with the receipt of information. Why haven't other people questioned this conceptual mixture?

Shinzuki: The reason is that game theory in the main stream was developed to treat ideal reasoning beings. The perfect memory assumption may be appropriate for such beings. Since game theorists have focused on games of perfect recall with FC, they managed to avoid (or ignore) such conceptual mixtures.

Since we are interested in the experiential foundation of knowledge or beliefs of a player with limited cognitive abilities including limited memory, we should care about such conceptual mixtures. As we eliminate FC and adopt its opposite NK, we should separate a player's memory ability from his "receipt of information," and treat both in a finite manner. The problem of memory is taken care of by a new concept, called a *memory function*.

[Shinzuki motions Jan to continue.]

Jan: Thank you for motivating the introduction of a memory function. We give a memory function $\mathfrak{m}_i^0$ to each player i, which describes the local (or temporal) memory at each node. The value, $\mathfrak{m}_i^0(x)$, of this function $\mathfrak{m}_i^0$ at node x is a sequence of information pieces received and actions taken previously that player i recalls at x. We call each value $m_i^0(x)$ a *memory thread*; the name thread is chosen since each memory may be used to weave a story. We allow a memory thread to be partial and even incorrect to capture forgetfulness of a player. In this way, we have separated the memory of a player from the physical phenomenon of receiving "information".

Shinzuki: By this separation, we widen the scope of game theory to include players with limited memories. It may help us to recall some examples for m_i^0. Could you give some for PL1 in Fig.5.1?

Jan: Okay, I start with an interesting example. It is the *recall*-1 memory function $\mathfrak{m}_i^{R1}$. When he reaches a node x, his temporal memory consists of the present information piece and his previous information piece + his action taken there. For example, the recall-1 memory function $\mathfrak{m}_1^{R1}$ of PL1 in Fig.5.1 is given as follows:

[Jan carefully writes formulae on the whiteboard.]

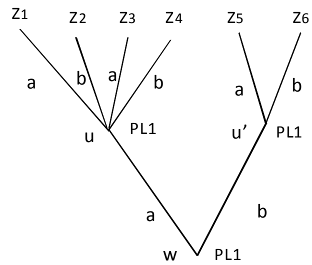

Figure 5.2: A Constructed View

$$\mathfrak{m}_1^{R1}(x_0) = \langle w \rangle,\ \mathfrak{m}_1^{R1}(x_2) = \mathfrak{m}_1^{R1}(x_3) = \langle (w,a), u \rangle, \mathfrak{m}_1^{R1}(x_4) = \langle (w,b), u' \rangle;$$

$$\mathfrak{m}_1^{R1}(z_1) = \langle (u,a), z_1 \rangle, \mathfrak{m}_1^{R1}(z_2) = \langle (u,b), z_2 \rangle, \mathfrak{m}_1^{R1}(z_3) = \langle (u,a), z_3 \rangle, \tag{9}$$

$$\mathfrak{m}_1^{R1}(z_4) = \langle (u,b), z_4 \rangle;\ \text{and } \mathfrak{m}_1^{R1}(z_5) = \langle (u',a), z_5 \rangle, \mathfrak{m}_1^{R1}(z_6) = \langle (u',b), z_6 \rangle.$$

Shinzuki: A lot of formulae. It is difficult to see the list.

[Jan is surprised and disgusted.]

Jan: You! You are at least partially responsible for such a complicated concept. Look at them carefully.

I'm quite pleased by the recall-1 memory. I use it often in practice. For example, when I try to recall how to get to some place, I remember it piece by piece. When I get to one corner, in my mind I can recall the next corner, and so on.

Shinzuki: Good. I'm starting to recall it piece by piece now. My brain works slowly as usual. Let me try to continue.

In (9), each value such as $\langle (w,a), u \rangle$ is called a memory thread, isn't it? Now, suppose hypothetically that this player, PL1, has all of those memory threads. Also, it is the basic assumption that he doesn't know his memory function $\mathfrak{m}_1^{R1}$, nor the game structure, that is, the NK assumption. Here we suppose that those threads are available for him. Then, consider the possibility that he connects those sequences like a jigsaw puzzle. Then, what is the result?

Jan: I can connect one with another if they share the same information pieces, i.e., $\langle w \rangle$ with $\langle (w,a), u \rangle$, and so on. The resulting picture from the list (9) is Fig.5.2.

Here, if this is his view, then it violates the standard theory of extensive games considerably. First, PL2 doesn't appear! Second, information pieces are used as connecting nodes. Third, at u, the same choice may lead to different outcomes.

Shinzuki: We call this process of recovering an extensive game from those memory threads *induction*. This already suggests a lot of new problems. But before going to them, it would be better to mention another example of a memory function. Suppose that PL1 can recall all of what he receives and what he has chosen. If we restrict his scope to his own nodes, then his memory function is called the *self-scope perfect-recall* function $\mathfrak{m}_1^{spr}$; here I give only two memory threads, since (9) took too much space:

$$\mathfrak{m}_1^{spr}(z_1) = \langle (w,a),(u,a),z_1\rangle \text{ and } \mathfrak{m}_1^{spr}(z_5) = \langle (w,b),(u',a),z_5\rangle. \tag{10}$$

Jan: When his scope includes PL2's information pieces and actions taken as well as PL2 himself, then this memory function is now called the *perfect-information* memory function $\mathfrak{m}_1^{pi}$. The first memory thread of (10) changes into

$$\mathfrak{m}_1^{pi}(z_1) = \langle \overset{1}{(w,a)}, \overset{2}{(v,a)}, \overset{1}{(u,a)}, \overset{\{1,2\}}{z_1}\rangle, \tag{11}$$

but the second of (10) remains the same. In this memory function, PL1 can peep at the information piece received and the action taken by PL2, which is the pair (v,a) in (11). In (11), we put the player's names who receive information pieces and move there. At the endpoint z_1, both players receive information pieces.

Remember, we even gave that complicated memory function called the "classical" memory function in [9]. It is too complicated to write here, but it showed that the standard theory with FC and perfect memory of each player can be expressed within the extended theory.

Shinzuki: Ok, let's focus on simple examples. From all the memory threads given by $\mathfrak{m}_1^{spr}$, the constructed view by PL1 is the same as Fig.5.2. From the memory threads given by $\mathfrak{m}_1^{pi}$, PL1 can reconstruct the view identical to Fig.5.1. However, he may not use nodes $x_0, x_1, ..., x_4$, since these nodes are not available to him; they are only available from our "objective" perspective, as theorists, to form a coherent story.

So, even with quite strong memory, a player recovers only some part from his accumulated memory threads. For a serious treatment of IGT, we can presume that a player experiences also only a small part of the game. I should emphasize that this partiality has two different sources: limited experiences and limited memory capabilities. After all, he cannot see the underlying nodes of the extensive game, which are hypothetical and only made from the objective observer's viewpoint.

Jan: Now, it is almost time for lunch and my stomach is starting to grumble.

Shinzuki: I'm still okay to continue, but it would be better to go to lunch now. Otherwise, it would be crowded in the cafeteria.

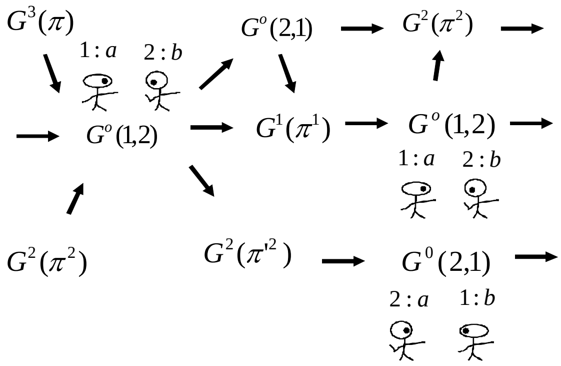

Figure 6.1: Social-web

6. Experiences in Recurrent Situations in the Social Web

Jan: The cafeteria was crowded with students. It is close to exam time! Argh..., I have a headache from anticipating grading so many exams, and complaining students afterwards.

Shinzuki: I'm sorry to hear it. I'm teaching only small classes, so I am less annoyed by grading. But if you teach well in a big class, then you must be popular with students. I envy you.

Jan: I think it would be better to leave envy aside and continue with more serious discussions.

Shinzuki: Okeydoke, Mr. Jan Hammer.

Consider how PL1 acquires the memory threads in (9). Above, we supposed that he has the full list of memory threads, without considering the source for them. It is critical that one play of the game might give the list of memory threads only along the path taken in the play. For example, if the players take the path to z_1 in Fig.5.1, then PL1 only receives the memory threads:

(12) $\mathfrak{m}_1^{R1}(x_0) = \langle w \rangle, \mathfrak{m}_1^{R1}(x_2) = \langle (w,a), u \rangle$ and $\mathfrak{m}_1^{R1}(z_1) = \langle (u,a), z_1 \rangle$.

To experience other memory threads, those players need to take a different path. Therefore, we assume a recurrent situation of this game.

Jan: We drew a nice picture in [9]. Let's see if I can recreate it on the whiteboard. We are interested in the social web surrounding one player. But for analytical purposes, we focus only on the occurrences of one particular game situation such as $G^0(1,2)$ in

Fig.6.1. So, the situation is simply described as:

$$\cdots \to G^0(1,2) \to G^0(1,2) \to \cdots$$

Then, other people have asked repeatedly: "Why don't you follow the repeated game approach?" I'm annoyed by this question. Even, I cannot use the word "repeated" or "repeatedly," which some game theorists relate automatically to the repeated game approach.

Shinzuki: Jan, my friend, they don't know any other way. They simply follow what they learned in game theory textbooks and/or classes.

Jan: I think so. The repeated game approach from standard textbooks formulates the entire repeated situation as a huge one-shot game. By this, which is a "sleight of hand," each player chooses his strategy before any interaction. Here, each must know the full description of the repeated situation, but this doesn't make sense for players who are learning the game structure. It only makes sense if we follow FC for the huge one-shot game.

Amazingly, this fact is still ignored by many game theorists and economists, and some of them are still working on repeated games even to describe learning and bounded rationality. We should ignore such work!

Shinzuki: Ha ha, now your anger turns to sarcasm. It would be more productive to move on.

Jan: OK. Where were we? Ah, yes, we started discussing the recurrent situation and why we don't follow the standard repeated game approach. The simple reason is that the players don't know the game structure, so they should not be expected to write a strategy for it. This raises the problem of how they play in such a situation.

We assume that they follow some regular behavior. In the beginning, the regular behavior is nothing more than some *default behavior*. Simply, a player chooses the first action when he receives an information piece and has to move. To make it all work, we assume that an information piece contains the description of available actions at that piece. In Fig.5.1, for example, each of the information pieces w, u, u', v contains the information that a and b are available actions there.

Shinzuki: Now, I can connect your description with (12). Suppose that they are always, or regularly, taking their default behavior a. Then, PL1's experiences are given by (12). We can connect those three memory threads, which is drawn in Fig.6.2. Notice that the figure has broken lines. The broken lines express the fact that PL1 knows those available actions but has no experiences of them or of their resulting outcomes. We call this type of view a "cane view" since it has only one line to the end, but each information piece has some broken branches going nowhere.

Jan: A new formulation of an extensive game must be able to handle this "cane view"

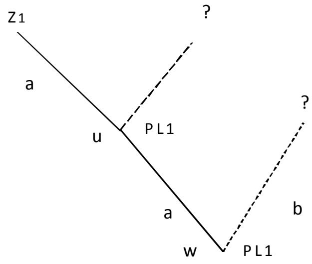

Figure 6.2: Cane View with Broken Branches

as well as standard ones. For such a purpose, we wanted directly to go to the new concept called an *information protocol*, forgetting an extensive game. But many of our colleagues have shown a conservative opinion by sticking to the traditional formulation of an extensive game, rather than accepting a new and direct formulation.

Shinzuki: But it turned out that their conservative opinion was somewhat beneficial for us.

Jan: That's why we gave two new formulations of an extensive game ([9], [10]);

(13): an extensive game in the strong and weak senses;

(14): an information protocol.

For (13), first we substitute information pieces for information sets. Still, the underlying tree structure is expressed by hypothetical concepts of nodes and edges. On the other hand, for (14), we forget such hypothetical concepts, and describe a target situation by historical causality of information pieces and actions taken. Then, we showed that these different constructs can be compared well, and are regarded as parallel theories.

Shinzuki: I remember our struggle well. Extensive games in (13) are classified into the strong and weak senses, and information protocols in (14) are classified by the choices of some conditions called non-basic axioms.

Jan: The formulation of an extensive game by Kuhn [17] can be regarded as (13) in the strong sense with information pieces substituted for information sets. To handle examples like Fig.5.2 and Fig.6.2, we introduced (13) in the weak sense. It was terrible to write down the whole definition of an extensive game, and also the definition in the weak sense appeared to be artificial; we didn't have enough confidence about the definitions in the weak sense, until we found the correspondences between (13) and (14).

Shinzuki: Indeed, a complete definition of an extensive game is horribly complicated. But this complication is not appreciated in our profession, since game theorists and economists merely use an extensive game to describe some economic situation. Their concern is the consideration of decision making in a given situation. They don't treat "extensive games" as objects of research but as a given environment.

For IGT, an extensive game is used both to be the objective description of the target game situation and a subjective description of a player's understanding of the situation. The former gives just an environment, but the latter becomes an object of a player's thinking, and thus an object of our research. Therefore, we needed to write the whole definition (13) as well as to introduce an extensive game in the weak sense.

Jan: However, the notion of an information protocol (14) was much easier than (13). This simplicity came because we didn't adopt the underlying tree structure in terms of a graph, since that part is hypothetical and non-tangible for a player. Starting with tangible information pieces and available actions, we directly described feasible sequences of pairs of information pieces and actions. Then, we gave concise axioms, that is, two *basic* and three *non-basic axioms*. Then we summarized the results in one paper ([10]).

Shinzuki: It is the most helpful finding that our axiomatization corresponds to the definition of an extensive game in the strong and weak senses. When all the axioms are assumed, an information protocol becomes equivalent to an extensive game in the strong sense; but when we choose some non-basic axioms, an information protocol is shown to be equivalent to an extensive game in the weak sense. Examples such as Fig.5.2 and Fig.6.2 can be captured as information protocols by appropriate choices of non-basic axioms.

In the beginning, we didn't think that we captured enough structures by an extensive game in the weak sense, since it seems quite simplistic for us. After all, the definition of an extensive game in the weak sense is stabilized by the correspondence between the weakened conditions for an extensive game and the non-basic axioms for an information protocol.

Jan: I know what you want to mention next. It must be their similarity to the correspondence between various axioms in epistemic logic and conditions in Kripke semantics. For example, Axiom T in (8) of the file of 8 years ago corresponds to the reflexivity of the accessibility relations in a Kripke frame (Hughes-Cresswell, [4], Kaneko [6]).

Shinzuki: Yes, you are right; I was surprised by that similarity. But, for our future research strategy on IGT as well as on epistemic logic, I think it would be much better to follow proof theory rather than Kripke semantics, especially, proof theory in the Gentzen-style, which has thus far been neglected. Kripke semantics together with proof theory are indispensable for our future developments.

Jan: Hey, man! I was pleased to find the similar correspondence results between epistemic logic and IGT. After that, you lost me. Once again, you left earth in search of some transcendental meta-level. I really don't know where you are heading now.

Shinzuki: I don't know either. I should simply follow the logical principles guiding us to further developments.

Jan: I may try to derive some sense from your connections later. At least they are now planted in my brain. Let's leave the topic now, and continue this discussion tomorrow morning. Today, I have to meet quite a few Ph.D students as the program director. Ah..., it is a headache.

Let's have pizza tomorrow for lunch in the restaurant we found near the beach avenue.

Shinzuki: Good idea! That pizza restaurant looks nice. See you tomorrow.

7. Experiences and Induction

[Jan came late to the visitor's office Shinzuki is using.]

Jan: Sorry I'm late. I couldn't get up this morning.

Shinzuki: Guten Morgen, Director Dr. Jan Hammer. You must be tired by meetings with graduate students yesterday. How were they?

Jan: Mm..., they talked a lot about causes for their problems and failures. I had headaches from listening to their worries. They didn't try to simplify their ways of thinking, but gave a lot of excuses. I told one student to practice simple and systematic routines with determination. I don't think he understood. By listening to so many problems, I was reminded of Tolstoy's claim: "*All happy families are a like; each unhappy family is unhappy in its own way.*"

Shinzuki: Good! I believe this metaphor applies to our developments in IGT too. We have various building blocks for game theory. Once we decide to follow the No-Knowledge Assumption (NK) instead of the Full-Cognizance Assumption (FC), we should follow the spirit of NK and modify each building block so as to be consistent with it. The spirit of NK is, more generally, related to the problems of "omniscience/omnipotence" in philosophical terms and of "bounded rationality" in economic terms. We should apply the spirit of NK to all building blocks simultaneously. If we make the mistake of limiting it to just one building block, then the resulting problem is just like an unhappy family in Tolstoy's "Anna Karenina".

Jan: Ah ..., yes, this is even related to our treatment of partial memory. Do you remember how we tried to treat imperfect recall by using information sets? We found that sometimes the combination of mixed strategies and a form of imperfect recall, called *a-loss recall*, cannot be distinguished from perfect recall[14]. I don't know if we can regard

[14] Kaneko-Kline [8].

this as a happy family, or an unhappy one, but at least we didn't capture our intention by just modifying one building block of an information set.

Shinzuki: Hmm...You are right. But keeping Tolstoy's theorem in mind, let's return to our discussions from yesterday.

We found we needed to weaken the concept of an extensive game, and to add an information protocol to handle problems of IGT. We discussed the concept of a memory function to separate the subjective memory of a player from the objective transmission of information described by an extensive game. What is next?

Aha...,we didn't yet discuss the accumulation of experiences. This is an integral part of IGT. Before a player can obtain an inductive derivation of an individual view, using either an extensive game (13) or an information protocol (14), he needs to accumulate his past experiences.

Jan: The accumulation of experiences is divided into two sub-processes. One is how new experiences are generated for a player, and the other is how experiences are changed into memories. As already mentioned yesterday, a player typically follows a pre-specified regular behavior pattern such as a default behavior. To accumulate some variety of experiences, some player needs to change his behavior, that is, he needs trials and errors. Then, he may have new experiences, caused by his own trials as well as by other players' trials.

Shinzuki: So he needs trials and errors, but not too much trials and errors. Do you remember how we found in Takeuchi *et al.* [27] that actual players, human subjects, in experiments could only remember a small number of experiences?

Jan: Yes, I do. This was even predicted in [9] by some epistemic postulates of forgetfulness and the need for repetition of the same memory to make it lasting. So we have some interesting finding that the trials are needed for a richer view. But, enough repetitions are needed to have a clear memory, not a broken line like in Fig.6.2. Here we see that we can capture our experimental observations only when we take the limitation of memory seriously and change the various building blocks accordingly.

Shinzuki: Very good. Let me continue with the theory of IGT. The memories that are experienced enough to become lasting form a *memory kit* for a player. The player uses this kit like a jigsaw puzzle box, treating each memory thread in the kit as a piece. For example, the memory threads given in (12) form a memory kit.

The next step is to combine the memory threads in the memory kit by some method like making the picture in a jigsaw puzzle. A constructed picture from the memory kit is called an *inductively derived view*. We may have multiple or possibly an infinite number of inductively derived views from a given memory kit. In this sense, the analogy to a jigsaw puzzle which is supposed to have a unique solution is inadequate.

Jan: Yes, it might be better to regard the memory kit as closer to a set of ingredients used by a master chef (or iron chef) to construct a meal. Some views are quite direct from

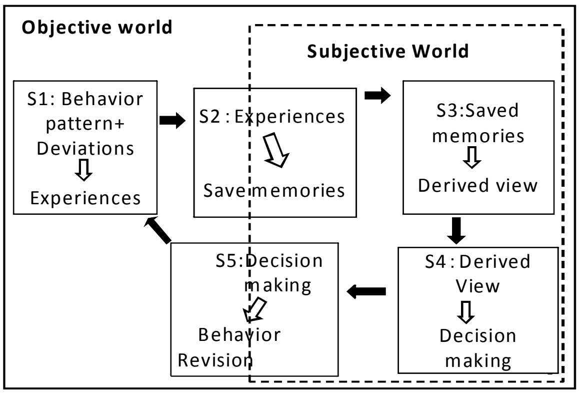

Figure 7.1: The Entire Process

the memory kit, while others are more indirect involving more elaborate combinations of the ingredients.

Shinzuki: The theory of extensive games is already too complicated for these purposes. The theory of information protocols given in [10] allowed us to make more progress in describing different types of induction. Anyhow, we have met so many alternative paths to pursue. Finally, we have written how many?,..., five or six papers, each of which is quite long. Ah..., we have worked a lot, but if we worked more, then we could dig deeper, but I had administrative meetings, teaching, watching TV, so time flew, Ah...

Jan: Hey Shinzuki! Don't get bogged down by retrospection. You start to sound like my graduate students. Our study and behavior should be systematic and determined.

Shinzuki: Ay, ay, sir, Mr. Hammer.

Jan: As the title of our first paper, *a basic scenario*, suggested, the entire process is long and winding. We discussed the generation and accumulation of memories in an informal manner, formulating basic premises as informal postulates in that paper.

You described this part as a process generating a memory kit, and then you discussed inductive processes and their possible results. What remains?

Shinzuki: Thank you for bringing me back to the point. Yes, we have another part. It describes how a player may use his inductively derived view for his decision making. The entire process is now described as Fig.7.1 from the basic-scenario paper.

Jan: Yes, we have completed one circle in the basic scenario like Fig.7.1. I suppose, in retrospect, we have done a lot.

Now, I'm getting hungry. Is it a good idea to have pizza in that restaurant in the beach avenue?

Shinzuki: I love it! It is slightly early, but let's celebrate for our progress in IGT with good pizza.

8. Toward a Bridge between Epistemic Logic and IGT

[They came back from the pizza restaurant in a sleepy mood.]

Jan: That pizza restaurant is nice. The young cook there is from Rome, so he makes good pizzas.

Shinzuki: Yes, he does, indeed. It would be nice to have that kind of restaurant in my town. Unfortunately, many nice restaurants have closed down now. We have still the sake-restaurant, Taishou. If it is closed, I should move somewhere else.

Jan: So, during lunch, you said we would still have something to discuss, didn't you? What else do we need to talk about now? Otherwise, I should take a small nap.

Shinzuki: Jan, do you remember the file of 8 years ago? We discussed problems of epistemic logic and game theory. Then, we have been working exclusively on IGT. Now, I think it is time to start to connect IGT with epistemic logic, not the time to have a nap.

Jan: Oh, let me see. In the file, we discussed "false beliefs", which may be a key for a connection between those theories. Inductively derived views almost surely include false elements, since experiences and one's memory ability are limited and partial. Um..., there must be a lot of elements for "false beliefs".

Shinzuki: You are right. Indeed, IGT is subject to "false beliefs" by its nature, in which sense it shares problems with epistemic logic. But we may still find a more direct and important connection.

Jan: Um ..., then where should we look? Do you have any clue?

Shinzuki: The clue must be "deduction"!

Jan: Do you say "deduction"? IGT is an inductive theory, isn't it?

Shinzuki: It may sound funny, but please think about the processes included in Fig.7.1. We have found various methods for even the main part, S3, of the figure, called *inductive derivation.* For example, combining memory threads from a memory kit, like a jigsaw puzzle, is a deductive method, indeed. We noticed the need for concatenation and other new procedures in [11] when we explored the recall-1 memory function.

Jan: I see. A method of induction can be regarded as a kind of an inference rule to connect one memory thread with another, by looking at information pieces, or whatever is needed. Do you mean this method can be formulated as some form of an inference rule or axioms?

Shinzuki: Yes, I do. It must be expressed as non-logical axioms, since we should avoid to mix the pure logical part with other parts.

Jan: Ok, then, the non-logical axioms determine an inductively derived view from a given memory kit, don't they?

[Shinzuki is now thinking.]

Shinzuki: But ..., I want to take one more step. Do you remember that we restricted our attention to the smallest view or minimal view for a given memory kit in [11] since otherwise we have too many?

Jan: Of course. So what? It is natural to choose a smaller view, since it requires us to think less.

Shinzuki: In IGT, we should take a lot of limitations into consideration. This means we are now departing from omniscience/omnipotence. Partiality corresponds to omniscience, but now we should seriously consider the problem of omnipotence.

To think about the smallest view may be regarded as one example for departure from omnipotence. Now, I should ask you to think about what the counterpart to smallness or limited potence is in epistemic logic. Given my penchant for proof theory, I suggest that we look into provability, which was written as $\vdash$ in the previous file from 8 years ago. In (1) of the previous file, we used it to express the existence of a proof of the statement.

$$(1) \qquad \vdash g^0, \mathrm{B}_1\mathrm{C}(g^1), \mathrm{B}_2\mathrm{C}(g^2) \to \mathrm{C}(\mathrm{Nash}(\mathbf{s}_{12}, \mathbf{s}_{22})).$$

Here, our criterion is whether or not there exists a proof of this statement. If we stop at provability or non-provability, we would still be under the assumption of omnipotence.

Jan: Now, I smell what you are up to. Before going to the heart of omnipotence, however, I want to ask about "common knowledge" included in the statement (1). To connect epistemic logic with IGT, we should drop the concept of common knowledge. For this aim, you have developed the theory of epistemic logics of shallow depths, haven't you?

Shinzuki: You are right. We should give limitations on interpersonal nesting structures of beliefs. We should first eliminate the concept of common knowledge from our logical system. The new logical systems, which we call *epistemic logics of shallow depths*, were already developed in several papers (cf., Kaneko-Suzuki [15], [16]).

Now it is the time to look into the provability $\vdash$. This must be the next step to move from omnipotence.

Jan: Shouldn't we be very careful about how we move from omnipotence? Otherwise, we may just create another unhappy family!

Shinzuki: We should take this step carefully and seriously. After incorporating limitations on *depths of interpersonal inferences* into epistemic logic, it would be natural to think about limitations at a fixed epistemic depth on *inferential abilities*. Let's do that too.

As already discussed, IGT involves a lot of bounded rationalities such as partial experiences, partial memories, various types of forgetfulness, perception bounds, etc. I used the philosophical term omniscience/omnipotence to discuss them.

Jan: In the economics literature, it is expressed as "bounded rationality" as a criticism of "super-rationality," isn't it? Herbert Simon [26] gave good descriptions of "bounded rationality," but he ended up with just one approach which has been called "satisficing behavior" with aspiration levels. Some people continue this approach blindly without realizing that it is just one approach to problems of bounded rationality.

Shinzuki: We have many more aspects to discuss as we have seen by developing IGT.

Jan: Could you give some concrete example to connect IGT to epistemic logic?

[Shinzuki stops talking and thinks for a quite some time.]

Shinzuki: Okay, let's consider one small example. Here we focus on the inferences within the mind of one player. Still, we have two different problems:

(15) shallow depths of interpersonal inferences; and

(16) limited inferential abilities at a given epistemic depth.

The former refers to whether or not i thinks about j's thinking. Shallow depths of interpersonal inferences mean that the nesting structure of "i thinks j thinks, ..." is limited, for example, even i does not think about j's thinking at all. On the other hand, limited inferential abilities could mean that j's thinking in the mind of player i is limited in terms of his (logical) calculation ability, or i's own calculation ability is limited. Both (15) and (16) break logical omnipotence, but in different ways.

For a concrete example of the above distinction, suppose that there are two stores in the market; PL1 is a *large* store, like a supermarket, and PL2 is a *small* store, a mini-mart. Consider the problem PL2 is facing and his interpretation of the situation.

Table 8.1

a	6000
b	5000
c	1000

Table 8.2

	s_1	s_2	s_3	s_4	s_5
a	5	6	7	8	9
b	15	16	17	18	19
c	(35)	(36)	(37)	38	39

PL2 has been in this situation with PL1 for quite some time. PL2 has accumulated experiences, and constructed his understanding of the situation. Suppose that his understanding is described by Tables 8.1 and 8.2. Table 8.1 gives PL2's understanding of

PL1's choices and payoffs. In PL2's mind, PL2's choice has only negligible influences on PL1's payoffs; and, indeed, PL2 regards PL1 as ignoring PL2's choice. Hence, PL1's payoff matrix in PL2's mind is determined only by the choice of PL1. Table 8.2 gives PL2's understanding of his own payoffs resulting from his own strategy s_t $(t = 1, ..., 5)$ and the choice a, b, or c by PL1.

Jan: Ok. So, PL1 in PL2's mind ignores PL2's choice. This may be natural as PL1 is large and significant, while PL2 is small and negligible. I make this judgement by looking at the relative size of PL1's and PL2's payoffs. But what do you mean by the parentheses in Table 8.2?

Shinzuki: The parentheses are just guesses he makes about the payoffs since, perhaps, he didn't experience them yet, or he has no recollection of those values.

Objectively speaking, there may be many other factors influencing the situation. But we suppose that those factors are small at least for PL2, and that he abstracts the situation as those two tables, ignoring the other factors. In this sense, the tables are a simple summary for PL2.

This is, more or less, the same as what we, theorists, are doing. Our theory must be an abstracted and simplified description of the real situation; if the theory is very accurate containing too many details, it could be too cumbersome and not understandable.

Jan: Perhaps, you have solved another question. Some people may be tempted to ask whether we need to assume that the situation is always the same for all people. This is also answered in the negative, isn't it?

Shinzuki: Yes. The subjective payoffs are simplified and abstracted. A lot of details are ignored. In the true situation, there are other factors influencing their payoffs. However, PL2 ignores such details as negligible factors.

Jan: Now, I can go to another question: How has PL2 acquired Table 8.1? If those players are in similar positions, that is, they are similar stores, then each could imagine the other's payoffs by putting his feet into the other's shoes. This argument may give the possibility of understanding other players' thinking, as well as some possible cooperation, which we have discussed in one of our papers [12].

But it seems we should not use this argument here, since the situation is quite asymmetric; the two players cannot substitute each other even in one player's mind. Where do we look for a source for PL2's knowledge?

Shinzuki: Um ..., perhaps, we should try to find good examples. A small store may get such knowledge from communication with other people as well as some books and consults, etc. It may be a good idea to do some empirical research related to this.

But here in (15) and (16) as mentioned before, we focus on limitations within the mind of one player. In [15], we used the term "interpersonal introspection" to distinguish these internal limitations of abilities from external limitations like communication. If we include communication and other external sources, we should limit them also.

Jan: I agree with you; we should be conscious of those phenomena, and then we may find a key for further research.

[Jan puts his hand on his forehead, and slowly starts to speak.]

By the way, we forgot the purpose of those tables, didn't we? We are supposed to discuss a possible bridge between IGT and epistemic logic, aren't we? What is your primary purpose of the above example?

Shinzuki: You are getting sharper and sharper, at least when you ask questions. I'm pleased.

I already wrote the two purposes of the above example, that is, (15) and (16). From the viewpoint of interpersonal inferences, the above example doesn't require the concept of common knowledge at all, since PL2's thinking about PL1's inferences is enough. In this sense, it is a problem of shallow epistemic depths within the mind of one player, PL2. Depth 2 is enough for PL2's decision making here.

Jan: In fact, if we eliminate those parentheses in Table 8.2, and if PL2 chooses a dominant strategy, i.e., s_5, then even the required epistemic depth is 1.

Shinzuki: That is right! Do you also notice that the required epistemic depths and decision criteria have some trade-off?

Jan: Trade-off? Suddenly, you become an economist! What is the trade-off you speak about?

Shinzuki: Be patient and I will show you. If PL2 adopts the *prediction-decision* criterion, then he needs to think about PL1's decision making, which means a greater depth of interpersonal inferences. On the other hand, if PL2 uses the *dominant-strategy* criterion, then his interpersonal inferences can be shallower, of depth 1 only[15]. Which do you think is more natural?

Jan: Obviously the second, since it requires us to think less.

Shinzuki: He he, you return to the standard thinking. Brace yourself, for I can give an argument that the first is simpler than the second.

The argument runs as follows. By using interpersonal inferences, PL2 can predict that PL1 will choose a. This prediction requires only two comparisons. Then, with this prediction, PL1 can choose s_5 by looking at the first row of Table 8.2 and making only four comparisons. In sum, he can reach his choice s_5 by $2 + 4 = 6$ comparisons.

On the other hand, if PL2 follows the dominant-strategy criterion, he can concentrate on Table 8.2, but he needs to think about all three rows of the table. The number of required comparisons to check s_5 to be the dominant strategy is $3 \times 4 = 12$. Hence, the second criterion requires twice as many comparisons as the first.

[15] See Kaneko-Suzuki [15] for a general treatment of prediction-decision criteria.

The pint is that predicting the choice of player 1 may simplify PL2's consideration for his choice; on the other hand, this includes interpersonal thinking.

Jan: I understood, but still, I'm not sure about where you want to go.

Shinzuki: Okay, let me try to explain. A key here, I think, is to have a clear-cut concept of "rationality" or "bounded rationality." "Rationality" must be regarded as an ideal limiting case of "bounded rationality", which must be the main target. Then, I would like to claim that the central theme is "bounded capability of logical inferences." This is a clear departure from the omnipotence assumption underlying provability.

Jan: Now, it becomes related to epistemic logic. It is good to hear that you identify "rationality" in terms of logical ability, rather than "utility maximization." In my opinion, "utility maximization" is just one element of a decision criterion.

Shinzuki: It is a small building block in game theory, and doesn't deserve the word "rationality."

Jan: Then, I should ask how "bounded capability of logical inferences" is related to the two decision procedures for Tables 8.1 and 8.2.

Shinzuki: Let me explain how to treat the above problem in epistemic logic. The interpersonal aspect of the problem is already quite clear. Decision making with the prediction-decision criterion requires the epistemic depth 2, while no interpersonal thinking is required for the dominant strategy criterion. This can be explicitly treated in the framework in [15].

Now, we take one step further, which must be a big leap. I calculated the numbers of payoff comparisons for the two decision criteria. These calculations are formulated in the framework in [15], too. But now, we will measure the number of inferences required for each claim. This can be done by counting the size of a proof of a claim in question. I calculated 6 and 12 comparisons for decision making with the two criteria. These are explicitly described by the number of inferences required for such statements. We should develop this procedure to count the number of inferences for a claim. This critically relies upon the Gentzen-style sequent formulation of the epistemic logics of shallow depths.

Jan: Um ..., this discussion becomes already quite long, and it may be endless. Can we eliminate some part? The Gentzen-style sequent calculus has the famous meta-theorem, called the *Cut-Elimination* theorem. Can you cut and eliminate some part of your discussion by it?

Shinzuki: The Cut-Elimination theorem is, indeed, important for the further advancement of our research. Hmm ..., but it must be opposite; actually, cut-elimination doesn't allow shortening the discourse, and it makes the discourse exhaustively explicit. Ah..., it is younger than that now being born in 1934/35[16]. Yet, it is still vividly alive and

[16] Gentzen [3]

central in proof theory.

Jan: Though it was born before my father, it is still very new.

Shinzuki: Yes, it is. Using the Gentzen-system, first, we need to construct a measure of deductive inferences. Then, we need to formulate the induction methods a player uses to have his view as deductive inferences. Then we can measure and compare the deductive inferences required for decision making and for the inductive inferences of having a view. This creates a bridge between the two theories.

In this manner, we can analyze a lot of problems related to "bounded rationality" appearing in Fig.7.1. Some problems force us to rethink the present thinking in epistemic logic such as a trade-off between deductive ability and memory.

Jan: Can you give a simple example of your "trade-off"?

Shinzuki: Japanese children are taught to memorize the multiplication table up to 9×9. Children of US and England memorize up to 12×12. So both groups can answer that $7 \times 8 = 56$ immediately in one step. However, if a Japanese boy is asked 11×12, he needs to calculate it by several steps. Which group do you belong to?

Jan: I was taught to memorize up to 12×12. However, sometimes I forget and use several steps to calculation. I guess I am a good and bad example. Does this have any implication for the present thinking in epistemic logic?

Shinzuki: Can you imagine how many steps are required to have $7 \times 8 = 56$ in the standard axiomatic system of arithmetic[17]? It is a tremendous number: Roughly, the required calculations are:

[Shinzuki now writes on the whiteboard]

$$
\begin{aligned}
7 \times 8 &= 7 \times (1 + (1 + (1 + (1 + (1 + (1 + (1 + 1))))))) \\
&= 7 + (7 \times (1 + (1 + (1 + (1 + (1 + (1 + 1))))))) \\
&= 7 + (7 + 7 \times (1 + (1 + (1 + (1 + (1 + 1)))))) \\
&= \ ...
\end{aligned} \tag{17}
$$

...by the distributive law. Mm ..., I'm not sure that distributivity is included as an axiom or is derived from more basic axioms. Let's skip this, and let me calculate a more basic part;

$$
\begin{aligned}
7 + 7 &= (1 + (1 + (1 + (1 + (1 + (1 + 1)))))) \\
&\quad + (1 + (1 + (1 + (1 + (1 + (1 + 1)))))) \\
&= ((1 + (1 + (1 + (1 + (1 + (1 + 1)))))) + 1) \\
&\quad +(1 + (1 + (1 + (1 + (1 + 1))))) \\
&= \ ...
\end{aligned} \tag{18}
$$

...by the associative law for summation $+$. Um ..., what is the next step?

[17] Cf., Mendelson [19], Chap.3.

[Jan in a small voice]

Jan: Ah,..., this guy drives me insane.

Stop, stop, please. Professor Kurai Shinzuki. Please do it in your notebook. You should state what you want to say in short sentences.

Shinzuki: Sorry about that, but my calculation is simpler than your list of memory threads in (9), isn't it? My point is that some criteria, such as independence, for an axiomatic system used in logic and mathematics should not be applied to the beliefs of a player. It suffices to memorize the results, and then a player isn't required to prove it each time. Mathematics is often taught with proofs. But nobody can reproduce a proof for every theorem that he knows.

Jan: This is good news. Now, I will be a person of memories rather than of inferences. I keep only theorems in my mind and can forget all proofs. I can drink beer now in a restaurant on the beach, as much I want, since my memory is all I need.

References

[1] Chellas, B., (1980), *Modal Logic*, Cambridge University Press, Cambridge.

[2] Fagin, R., J.Y. Halpern, Y. Moses and M. Y. Verdi, (1995), *Reasoning about Knowledge*, The MIT Press, Cambridge.

[3] Gentzen, G., (1934/35), Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift* 39, 176–210, 405–431. English translation: Investigations into logical deduction. In: Szabo, M., (ed.) *The Collected Papers of Gerhard Gentzen.* North-Holland (1969), Amsterdam:

[4] Hughes, G. E. and Cresswell, M. J., *A Companion to modal logic.* London: Methuen &Co. (1984).

[5] Kaneko, M., (1999), Epistemic considerations of decision making in games. *Mathematical Social Sciences* 38, 105–137.

[6] Kaneko, M., (2002), Epistemic Logics and Game Theory: Introduction, *Economic Theory* 9, 7-62.

[7] Kaneko, M., (2004), *Game Theory and Mutual Misunderstanding*, Springer, Berlin.

[8] Kaneko, M., and J. J. Kline, (1995), Behavior strategies, mixed strategies and perfect recall, *International Journal of Game Theory* 24, 127–145.

[9] Kaneko, M., and J. J. Kline, (2008a), Inductive Game Theory: A Basic Scenario, *Journal of Mathematical Economics* 44, 1332–1263.

[10] Kaneko, M., and J. J. Kline, (2008b), Information Protocols and Extensive Games in Inductive Game Theory, *Game Theory and Applications* 13, 57-83.

[11] Kaneko, M., and J. J. Kline, (2008c),Partial Memories, Inductively Derived Views, and their Interactions with Behavior, to appear in *Economic Theory.*

[12] Kaneko, M., and J. J. Kline, (2009), Transpersonal Understanding through Social Roles, and Emergence of Cooperation,

http://www.sk.tsukuba.ac.jp/SSM/libraries/pdf1226/1228.pdf

[13] Kaneko, M., and A. Matsui, (1999), Inductive Game Theory: Discrimination and Prejudices, *Journal of Public Economic Theory* 1, 101–137.

[14] Kaneko, M., T. Nagashima, N.-Y. Suzuki, and Y. Tanaka, (2002), A Map of common knowledge logics, *Studia Logica* 71, 57-86.

[15] Kaneko, M., and N-Y. Suzuki, (2002), Bounded Interpersonal Inferences and Decision Making, *Economic Theory* 9, 63-104.

[16] Kaneko, M., and N-Y. Suzuki, (2003), Epistemic models of shallow depths and decision making in games: Horticulture, *Journal of Symbolic Logic* 68, 163-186.

[17] Kuhn, H. K., (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.

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

[19] Mendelson, E., (1987), Introduction to mathematical logic. Monterey: Wadsworth.

[20] Moser, P. K., (1986), ed., *Empirical Knowledge: Readings in Contemporary Epistemology.* Rowman & Littlefield Publishers, Inc. Chicago.

[21] Meyer, J.-J. Ch. and van der Hoek, W., (1995), *Epistemic logic for AI and computer science.* Cambridge: Cambridge University Press.

[22] Nash, J. F., (1951), Noncooperative Games, *Annals of Mathematics* 54, 286–295.

[23] Plato, (1941), *The Republic of Plato*, Translated by F. M Cornford. London: Oxford University Press.

[24] Savage, L., (1954), *Foundations for Statistics,* Dover Publication Inc. New York.

[25] Schulte, O., (2002), Minimal Theory Change and the Pareto Principle, *Economic Theory* 9, 105-144.

[26] Simon, H., (1955), A Behavioral Model of Rational Choice, *Quarterly Journal of Economics* 69, 99-118.

[27] Takeuchi, A., Y. Funaki, M. Kaneko, J. J. Kline, (2010), An Experimental Study from the Perspective of Inductive Game Theory, coming soon.

[28] Terutoshi, Y. Okitsu K. and Enomoto, S., (eds.), (1980), *Collections of meiji-taisho comic stories* (in Japanese). Vol. 3. Tokyo: Kodansha

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