KNOWLEDGE TRANSFER AS A RATIONAL CHOICE: A DECISION THEORETIC CHARACTERIZATION

Yasuo Sasaki
School of Knowledge Science, Japan Advanced Institute of Science and Technology
1-1, Asahidai, Nomi, Ishikawa, Japan

ABSTRACT

This paper presents a game theoretical framework to analyze the possibility of knowledge transfer about a game structure. We study, under asymmetric awareness of two agents in a game, what kind of knowledge can, or cannot, be transferred from one agent who has more knowledge to the other agent with restricted knowledge prior to playing the game. Such situations are generally characterized as a particular class of a recently developed framework of extensive-form games with unawareness and analyzed with the solution concept called rationalizability. We show some properties of such games and example analysis. Also we discuss some ideas to incorporate, in a more general way, inferences of the agents about knowledge transfer under asymmetric awareness.

Keywords: Knowledge transfer, Game theory, Rationalizability, Awareness.

INTRODUCTION

If I know something of which you are not aware and it is relevant to the game we play, then it can be my important decision problem whether I should tell about it to you. If I am a rational decision maker, I would decide if I transfer the knowledge to you taking into account the impact of the decision, that is, how the knowledge transfer can change your perception about the game and thus your behavior and how, as a result, such changes can affect the level of my benefit in the game. The present study aims at providing a formal framework to analyze the possibility of knowledge transfer in this kind of situation by regarding it as a rational choice of an agent (decision maker).

We use game theory because it is probably the most developed formal framework for interactive decision-making. The standard game theory, however, is not sufficient for the purpose because it usually assumes common knowledge of the game structure. In other words, the agents are typically assumed to possess knowledge about it as much as the modeler does (Myerson, 1991). Although many interesting applications of communications among the agents under asymmetric information such as signaling (Spence, 1973) and cheap-talk (Crawford and Sobel, 1982) have been studied mainly in economics, they analyze the agents' rational behaviors in a game with common knowledge: the signal or message exchanged there does not alter one's perception about the game structure itself. In another line of game theoretical approach to knowledge transfer, there have been several studies how, in the context of knowledge management (e.g. Nonaka and Takeuchi, 1995), people can collaborate in creating and sharing knowledge (Samaddar and Kadiyala, 2006; Bandyopadhyay and Pathak, 2007; Li and

Jhang-Li, 2010). They regard knowledge as a sort of public goods and thus the problems essentially fall into public goods provisions.

On the other hand, recently a game theoretical framework that deals with possibly unaware agents (e.g. Heifetz et al., 2013; Halpern and Rêgo, 2014; Schipper, 2014). Games with unawareness are distinguished from well-known games with incomplete information (Harsanyi, 1967) in that, in the former, the agents may not be able to even conceive of some relevant components of the game. (Hypergame theory, another game theoretical framework that deals with misperceptions of agents, can also formalize similar situations (Bennett, 1977; Wang et al., 1988; Sasaki and Kijima, 2012)). In this paper, by using this framework and its solution concept, we formulate and analyze the problem of knowledge transfer under asymmetric awareness in games. While there are many types of asymmetric awareness which games with unawareness can formulate, we only study a simple form of it in two-agent games: one agent knows the true game while the other is unaware of some component(s) of it, and the latter agent's unawareness is known by the former agent. Then we analyze what kind of knowledge can, or cannot, be transferred from the agent who has more knowledge to the other agent with restricted knowledge prior to playing the game.

This paper is organized as follows. Following the introduction, we first characterize such situations as a particular class of extensive-form games with unawareness, which we call knowledge transfer games. We also introduce a solution concept called rationalizability (in games with unawareness) to analyze them and present some general properties. Then we show an example analysis. After that, since we admit that the framework presented here is just the first step to study the problem of knowledge transfer under asymmetric awareness, we discuss several ideas to extend our model. All the proofs of mathematical propositions will be given in Appendix.

ANALYTICAL FRAMEWORK

The Problem Setting: Knowledge Transfer Games

Consider a two-agent normal-form game played by agent 1 (she) and 2 (he). Suppose agent 1 perceives the game correctly while agent 2 is not aware of some action(s) of agent 1 and/or 2 himself. That is, he views a restricted game, and we suppose that he believes the restricted game is common knowledge. Agent 1 knows such a view of agent 2, that is, she believes that he perceives the restricted game and believes it is common knowledge. In this sense, she is confident that she has more knowledge about the game (i.e. available actions of them) than him. Therefore it becomes her considerable choice prior to playing the game whether she tells him such knowledge. On the other hand, from agent 2's point of view, he just believes they are playing a standard normal-form game with common knowledge, thus he would not have such a motivation of knowledge transfer (and indeed he cannot do that). (If we consider the possibility of cheating, this may not be the case. We will discuss this issue later.) Therefore, we only consider the possibility of knowledge transfer from agent 1 to 2. For this purpose, before playing the normal-form game, we add her choice of whether she tells something to her opponent, and if she does, what kind of knowledge she is going to tell him. Such a knowledge transfer, if done, will make agent 2 aware of the new knowledge, i.e. some action(s) of which he was previously unaware. In this sense, it can alter his view about the game, and hence his behavior. Agent 1 makes a decision taking into account such a possible change in her opponent's perception led by her knowledge transfer.

Formally, such a situation can be modeled as a particular class of extensive-form games with unawareness (Heifetz et al., 2013; Halpern and Rêgo, 2014). Since the general framework is too complicated to be presented here, we formalize this class of game with unawareness as a "knowledge transfer game" (KT game). Let us start with the objective description of the game faced by the two agents. This is a normal-form game played by agent 1 and 2, $G^0 = \{A_1^0, A_2^0; u_1^0, u_2^0\}$, where A_i is agent i's non-empty finite action set and $u_i^0 : A_1^0 \times A_2^0 \to \mathbb{R}$ is i's real-valued utility function for $i \in \{1,2\}$. As a technical setting, we require at least one agent's action set contains two or more actions so that, in a restricted game perceived by (unaware) agent 2, the both agents have non-empty action sets.

Then, a KT game based on the game G^0 is defined as an extensive-form game with unawareness as depicted in Figure 1. In the figure, there are *n* normal-form games, G^1, …, G^n. For every k, $G^k = \{A_1^k, A_2^k; u_1^k, u_2^k\}$, where, for any $i \in \{1,2\}$, $A_i^k (\neq \phi) \subseteq A_i^0$ and $u_i^k : A_1^k \times A_2^k \to \mathbb{R}$. First, let us see G^1 in the figure. This is a normal-form game perceived by agent 1 when she tells her opponent nothing, which is supposed to be identical to G^0. (In order to allow for simultaneous moves in an extensive-form game, we follow Osborne and Rubinstein (1994, Ch. 6).) At this point, agent 1 believes agent 2 perceives not G^1 but a restricted game, G^2, where $A_1^2 \subset A_1^0$ and/or $A_2^2 \subset A_2^0$. That is, she considers he is unaware of $A_1^2 \setminus A_1^0$ and $A_2^2 \setminus A_2^0$. (At least one of these is non-empty.) In G^2, we assume $u_i^2(a) = u_i^0(a)$ for any $a \in A_1^2 \times A_2^2$.

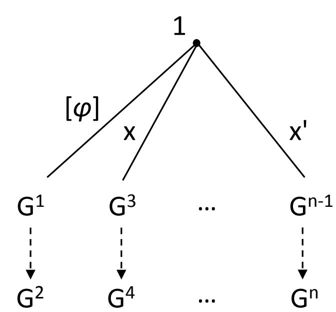

Figure 1. The Structure of Knowledge Transfer Game

In the beginning of the KT game, she chooses what kind of knowledge of which her opponent is unaware she tells him. She can also choose telling him nothing. Denote the first decision node by h_0 and the set of her choices there by $\bar{A}_1$. Then, $\bar{A}_1 \subseteq 2^{A_1^2 \setminus A_1^0} \times 2^{A_2^2 \setminus A_2^0}$ and we assume (ϕ, ϕ), which we denote $[\phi]$ hereafter, and at least one more element is included in $\bar{A}_1$. Agent 1's choice of $x \in \bar{A}_1$ means that she tells x to agent 2. For example, if $x = (\{a_1\}, \{a_2\})$ with $a_1 \in A_1^2 \setminus A_1^0$ and $a_2 \in A_2^2 \setminus A_2^0$, this is interpreted as her telling him, "I have a_1 and you have a_2 (of which you are unaware) in the game we play." In particular, her choice of $[\phi]$ means she tells nothing to him. For any $x \in \bar{A}_1$, when x leads to a game G^k, we write $h(x) = G^k$, which is interpreted as, when she chooses x at h_0, then her perception about the game is G^k. Denote the set of these normal-form games that can be given as $h(x)$ for some $x \in \bar{A}_1$ by Γ^1 (i.e. every G^k with odd-numbered k in Figure 1). Here we assume agent 1's knowledge transfer may entail some cost, that is, if she chooses any option except for $[\phi]$ at h_0, she has to incur a non-negative cost, c (in her utility term). Thus every G^k except for G^1 that can be reached by some $x \in \bar{A}_1$ has the same game structure as G^0 except that $-c$ is added to agent 1's utility value at every outcome.

Each G^k in Γ^1 describes agent 1's perception about the normal-form game they play. As stated above, she understands the game is G^0 and, when she tells her opponent something, it will take the cost c. On the other hand, at each game reached, her view about agent 2's view can be largely different. This depends on her choice at the beginning, i.e., which knowledge can be transferred. We describe her perception about her opponent's view with an injective function f: the fact that she believes he perceives a game G^l, and indeed he perceives it, when she chooses $x \in \bar{A}_1$ at h_0 is described as $f(h(x)) = G^l$. For instance, $f(h([\phi])) = f(G^1) = G^2$. The dotted arrows in the figure express these mappings. In general, in a game G^l that can be given as $f(h(x))$ with $x = (x_1, x_2) \in \bar{A}_1$, $A_i^l = A_i^2 \cup x_i$ for any $i \in \{1,2\}$. Thus, we assume that, whenever she tells him, "I have an action a_1 (of which you are unaware)," she believes he will add a_1 to the action set of agent 1, and indeed he does so. The same thing goes whenever she refers to some action(s) of agent 2. (Thus we assume her message is always "credible." For this assumption, see the discussion later.) Therefore, any two normal-form games obtained by a mapping by f have different action sets with one another. Denote the set of all the candidates of agent 2's perception, namely all the normal-form games that can be given as $f(h(x))$ with some $x \in \bar{A}_1$, by Γ^2 (i.e. every G^l with even-numbered l in Figure 1). Furthermore we assume that her knowledge transfer cost is common knowledge. Thus, in every G^l except for G^2 in Γ^2, $u_1^l(a) = u_1^0(a) - c$ and $u_2^l(a) = u_2^0(a)$ for any $a \in A_1^l \times A_2^l$. Consequently, the number of the normal-form games in the KT game, n is $2|\bar{A}_1|$.

In a game with unawareness, unlike standard games, an agent may not know the whole structure of the game. In a KT game, when agent 1 chooses $x \in \bar{A}_1$ and $f(h(x)) = G^l$, agent 2 just considers they are playing a standard game with common knowledge, $G^l \in \Gamma^2$. On the other hand, agent 1 is supposed to possess a view about the whole game structure in the same way as the modeler: she can view the whole picture of Figure 1. The perceptional difference is important when considering each agent's decision making in the KT game. Agent 1 would make a choice at h_0, taking into account the impact of her choice on agent 2's possible perceptional change and his behavior in light of his perception, and then act rationally in the two-agent normal-form game. On the other hand, agent 2, in his point of view, would just make a choice in a standard normal-form game, and agent 1 knows this.

Solution of the Game: Rationalizability

As a solution concept of the game with unawareness, we employ (extensive-form) rationalizability defined for extensive-form games with unawareness by Heifetz et al. (2013). Since the general definition of it is somewhat complex, we here define rationalizability in a KT game presented above. In Appendix, we show that our definition can surely be captured by their original definition.

The concept of rationalizability was originally introduced in game theory by Pearce (1984). In a standard game, an agent's rationalizable action is considered as such an action that can be taken when the game structure and the agents' rationality are common knowledge. It is a weaker concept than Nash equilibrium since Nash equilibrium requires conjectures of the agents (about their choices) to be consistent while rationalizability does not. Hence, rationalizability in an extensive-form game with unawareness implies possible outcomes of the game when the agents' rationality is common knowledge, the agents' perceptions about the game structure is given as the definition of the game with unawareness and the play is one-shot. (For a more rigorous characterization of rationalizability, see Battigalli (1997); Perea (2012).)

First, we define rationalizability for normal-form games. Consider a generic two-agent normal form game, $G = \{A_1, A_2; u_1, u_2\}$. For each agent i ($\in \{1,2\}$), let ΔA_i be the set of probability distributions on A_i. Hence $\delta_i \in \Delta A_i$ can be interpreted as i's mixed action. Let i's opponent be j. We denote agent i's expected utility when i takes $a_i \in A_i$ and j takes $\delta_j \in \Delta A_j$ by $Eu_i(a_i, \delta_j)$. Also let $H_i(0) = A_i$ and, for any positive integer t, $H_i(t) = \{a_i \in A_i | [\exists \delta_j \in \Delta H_j(t-1)][\forall a_i' \in H_i(t-1), Eu_i(a_i, \delta_j) \geq Eu_i(a_i', \delta_j)]\}$. Then $a_i^* \in \bigcap_{t=0}^{\infty} H_i(t)$ is called agent i's rationalizable action. It is such an action of agent i that can survive iterated eliminations of actions that cannot be a best response to the opponent's choice under any belief of i about j's choice. Let us denote the set of i's rationalizable actions in a game G by $R_i(G)$, which is a subset of A_i. It is known that, in general, $R_i(G)$ is non-empty. Hereafter, in a KT game, we also denote agent i's expected utility function in a normal-form game G^k defined as above by Eu_i^k.

With respect to (extensive-form) rationalizability in a KT game, if agent 2 makes a decision according to it, then he must choose an action in $R_2(G^l)$ in every $G^l \in \Gamma^2$. Then, knowing that, agent 1 makes a decision rationally. Since she recognizes the extensive-form game in which she acts multiple times, her rationalizability is defined on her strategies. Although, in general, an agent strategy is defined as a combination of her actions at every information set, we here define it simply as a combination of her choices at h_0 and in the normal-form game led by the first choice there since choices in the other normal-form games are irrelevant when we consider the solution concept (see Appendix). That is, $s_1 = (x, a_1)$ is her strategy with $x \in \bar{A}_1$ and $a_1 \in A_1^k$ such that $h(x) = G^k$. Denote the set of her strategies by S_1.

Now agent 1 has a belief about how agent 2 would act in each game in Γ^2. This is described as probability distributions on his (normal-form) rationalizable actions in these games, namely an element of $\prod_{G^l \in \Gamma^2} \Delta R_2(G^l)$. Denote this set of her beliefs by Σ_2.

Agent 1's strategy $s_1 = (x, a_1) \in S_1$ under a belief $\sigma_2 \in \Sigma_2$ gives her expected utility $Eu_1(s_1, \sigma_2)$, which is calculated as $Eu_i^k(a_1, \delta_2)$ when $h(x) = G^k$, where δ_2 is agent 2's mixed action in G^k, which is specified by the probability distribution on $R_2(f(G^k))$ in σ_2. Then we call a strategy of agent 1 is rationalizable when it is a best response for her under some belief. Formally, $s_1^* \in S_1$ is her rationalizable strategy if and only if there exists $\sigma_2 \in \Sigma_2$ such that $Eu_1(s_1^*, \sigma_2) \geq Eu_1(s_1, \sigma_2)$ for any $s_1 \in S_1$. Denote the set of her (extensive-form) rationalizable strategy in the KT game by R_1, which is a subset of S_1. Since it is known that each agent has at least one rationalizable strategy in an extensive-form game with unawareness (Heifetz et al., 2013), we have the following remark on the existence of rationalizable strategies in any KT games.

Remark 1: In a KT game, $R_1 \neq \phi$ and $R_2(G^l) \neq \phi$ for every $G^l \in \Gamma^2$.

Properties of Knowledge Transfer Games

With respect to the role of knowledge transfer cost, we have the following propositions.

Proposition 2: In a KT game, if $x^* = [\phi]$ for any $(x^*, a_1^*) \in R_1$ when the knowledge transfer cost is c, then the same thing holds when it increases to any $c'(> c)$.

Proposition 3: In a KT game, if there exists $x^* \neq [\phi]$ such that $(x^*, a_1^*) \in R_1$ for some $a_1^* \in A_1^k$, where $h(x^*) = G^k$, when the knowledge transfer cost is c, then $(x^*, a_1^*) \in R_1$ when it decreases to any $c'(< c)$.

The two propositions are quite intuitive. Proposition 2 claims that, when any kinds of knowledge transfer cannot occur in some situation, the same thing holds if its cost increases. Likewise, Proposition 3 says that when some kind of knowledge transfer can occur in some situation, it is also possible if its cost decreases. Generally, based on the definition of rationalizability, the condition of knowledge transfer can be restated as follows.

Proposition 4: In a KT game, for $x^* \in \bar{A}_1$ such that $h(x^*) = G^k$, there exists $a_1^* \in A_1^k$ such that $(x^*, a_1^*) \in R_1$ if and only if, for any $G^{k'} \in \Gamma^1 \setminus \{G^k\}$,

$$\max_{\delta_2 \in \Delta R_2(f(G^k))} \left(\max_{a_1 \in A_1^k} Eu_1^k(a_1, \delta_2) \right) \geq \min_{\delta_2 \in \Delta R_2(f(G^{k'}))} \left(\max_{a_1 \in A_1^{k'}} Eu_1^{k'}(a_1, \delta_2) \right).$$

That is, it is rationalizable for agent 1 to choose $x \in \bar{A}_1$ at h_0 if and only if, given that agent 2 will choose his rationalizable action in each $G^l \in \Gamma^2$, the maximum value of her expected utility when she takes a best response in $h(x)$ under any belief is higher than or at least equal to the minimum value of that in any normal-form games in Γ^1 other than $h(x)$. Thus, based on the statement, we can check if a transfer of particular knowledge $x \in \bar{A}_1$ can take place. (If $x = [\phi]$, we can check if not telling anything can be rationalizable.) Then the next proposition follows Proposition 4 straightforwardly. It states that if agent 1 makes a decision at h_0 according to the max-min principle (given that her opponent will always choose his rationalizable action), the choice is always rationalizable. That is, it is rational for her to make a decision at first so that she maximizes the minimum value of her expected utility when she takes a best response under any belief in the subsequent normal-form game.

Proposition 5: In a KT game, for $x^* \in \bar{A}_1$ such that $h(x^*) = G^k$, there exists $a_1^* \in A_1^k$ such that $(x^*, a_1^*) \in R_1$ if, for any $G^{k'} \in \Gamma^1 \setminus \{G^k\}$,

$$\min_{\delta_2 \in \Delta R_2(f(G^k))} \left(\max_{a_1 \in A_1^k} Eu_1^k(a_1, \delta_2) \right) \geq \min_{\delta_2 \in \Delta R_2(f(G^{k'}))} \left(\max_{a_1 \in A_1^{k'}} Eu_1^{k'}(a_1, \delta_2) \right).$$

EXAMPLE ANALYSIS

We use a variant of Bach-Stravinsky-Mozart game studied in Heifetz et al. (2013), which is an extension of so-called battle-of-the-sexes game. Consider the following situation, which we formulate as a KT game of Figure 2. Two agents, 1 (she) and 2 (he), independently make decisions about which concert they go to. There are three concerts, Bach (B), Stravinsky (S) and Mozart (M), and their utilities are illustrated in the upper left matrix (G^1). If they both go to the Mozart concert, it is a Nash equilibrium which is Pareto-optimal outcome of the game, but now suppose only agent 1 is aware of Mozart while agent 2 is unaware of it and considers there will be only two concerts, Bach and Stravinsky. Agent 1 knows such a restricted perception of agent 2, which is described as the lower left matrix (G^2). Thus if she does not tell about the Mozart concert, which we denote $[\phi]$, she views the game is G^1 while she believes he views G^2. Before their choices of concerts, she can also tell him the existence of the Mozart concert, which we denote $[M]$, with cost $c \geq 0$. Thus, when she does so, her perception of the game becomes the upper right matrix (G^3). In this case, agent 2's view contains all the three concerts and reflects the knowledge transfer cost of agent 1, therefore it can be expressed as the lower right matrix (G^4).

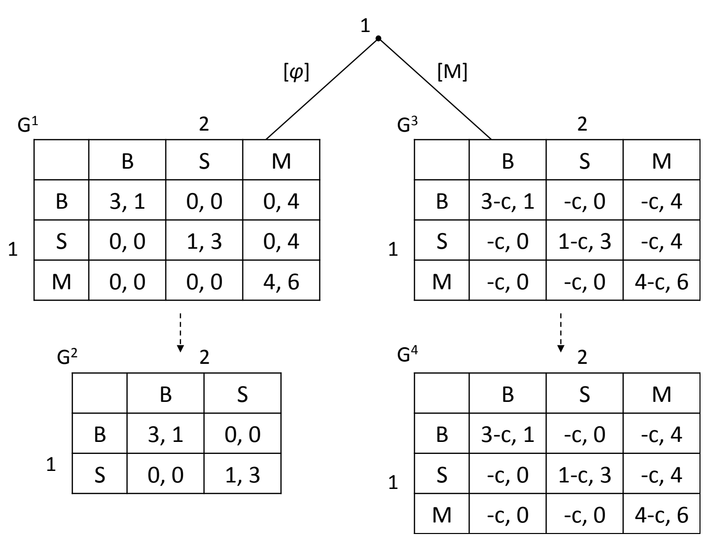

Figure 2. Bach-Stravinsky-Mozart Game

In this KT game, $R_2(G^2) = \{B, S\}$ and $R_2(G^4) = \{M\}$, hence, according to rationalizability, agent 2 choose these actions. In this case, his choice is irrelevant to agent 1's knowledge transfer cost because going to the Mozart concert is his dominant action in G^4. On the other hand, agent 1's rationalizable strategy depends on the cost and the set of it is calculated as follows. First, when $c < 1$, $R_1 = \{([M], M)\}$. In this case, since agent 1 has probability-one belief that agent 2 would choose Mozart in G^3, it is her dominant strategy to tell about the Mozart concert first and choose it. Second, when $1 \leq c < 3$, $R_1 = \{([\phi], B), ([M], M)\}$. In this case, depending on her belief on $R_2(G^2)$, there are two rationalizable strategies of agent 1: not telling about the Mozart concert and going to Bach, and telling it and going to Mozart. The former is rational when she believes her opponent will choose Bach with at least $(4 - c)/3$ probability in G^1, while the latter is rational when she believes her opponent will choose Stravinsky with at least $(c - 1)/3$ probability in G^1. Third, when $3 \leq c \leq 10/3$, $R_1 = \{([\phi], B), ([\phi], S), ([M], M)\}$. In this case, in addition to the two strategies in the previous case, not telling about the Mozart concert and going to Stravinsky can also be rationalizable under some belief. Finally, when $10/3 < c$, $R_1 = \{([\phi], B), ([\phi], S)\}$. In this case, under any belief, telling about the Mozart concert cannot be a part of a rationalizable strategy. Although she perceives G^1, her inference is essentially same as that in a standard battle-of-the-sexes game.

To sum up, from the viewpoint of knowledge transfer, when its costs is sufficiently low, only telling about the Mozart concert is rationalizable, while when the cost is sufficiently high, only not telling can be rationalizable. There is also a middle range of it, where both telling and not telling can be rationalizable. Thus Proposition 2 and 3 can be confirmed by these facts. Furthermore, in the middle range, the set of agent 1's possible choices when she chooses not telling is different: when the cost is relatively low, she only chooses Bach in G^1 (i.e. the second case above), while when it is relatively high, she may choose both Bach and Stravinsky there (i.e. the third case above). As for the implications of Proposition 4 and 5, in G^1, the maximum and minimum values of agent 1's expected utility when she takes a best response under any belief are 3 and $3/4$, respectively, while both the maximum and the minimum are $4 - c$ in G^3. Based on these values, the reader can see the propositions hold here.

DISCUSSIONS

The framework of KT games introduced in this paper can be extended in several ways. We consider the model and analysis presented above are just the first step to study knowledge transfer under asymmetric awareness in games. This section shows some ideas of possible extensions as the concluding remarks of this paper.

Credibility of Knowledge Transfer and the Possibility of Cheating

The framework presented above assumed that whenever agent 1 tells some knowledge to agent 2, he believes her message is true: it is always credible for him. It, however, may not be reasonable in some situations. Consider the following example, which is a variant of the Bach-Stravinsky-Mozart game above. Now the objective situation is the left matrix in Figure 3, which is also agent 1's view. The Mozart concert is only available to her. Agent 2 is unaware of it and just sees the standard battle-of-the-sexes game, the right matrix, and agent 1 knows his unawareness of Mozart. Then it is her choice whether or not she tells about it to him. The problem here is the credibility of the message when she does so. If she tells him, "I have another option of which you are not aware, the Mozart concert," how would he interpret this?

1 \ 2	B	S
B	3, 1	0, 0
S	0, 0	1, 3
M	2, 0	2, 0

1 \ 2	B	S
B	3, 1	0, 0
S	0, 0	1, 3

Figure 3. A Variant of the Bach-Stravinsky-Mozart Game

Our answer is that it is far from clear whether or not he will update his view from the right matrix to the left one. In the left game, Mozart strictly dominates Stravinsky for agent 1, thus she would not choose Stravinsky. Given this perspective (i.e. if we eliminate agent 1's Stravinsky), going to the Stravinsky concert is weakly dominated for agent 2. Hence it would be more plausible that he will choose Bach (though Stravinsky is also rationalizable), then agent 1 will be able to achieve the best outcome for her by choosing Bach as well. Based on this kind of inference, agent 2 might consider that his opponent is possibly cheating, that is, the Mozart concert is actually not available to her, so that she will get the best payoff for her. Thus the credibility of such a message is non-trivial. (Feinberg (2008) also discusses this issue with a similar example.) Such a problem can often be seen in the real world as well. For example, in a war between two countries, what if one of them declares, "We now have a new weapon"? It would not be clear if the message can immediately be accepted by the other side.

Furthermore, based on this example, we can think of another problem, the possibility of agent 1's strategic cheating. Now suppose the objective situation is just the standard battle-of-the-sexes game, and agent 1 understands this: the Mozart concert is indeed not available to her. But if she successfully convinces agent 2 to believe the game is the left matrix of Figure 3, she can avoid the coordination problem and achieve the best result for her. Though it is not clear that telling such a lie, namely saying, "I have actually one more option," is credible for her opponent as discussed above, it might become her considerable choice. (Hämäläinen (1981) discusses the possibility of cheating of the leader in a Stackelberg game, though it is not cheating on the game structure itself.)

Scope of Agent 2's Inference

We assumed that agent 2 believes he just plays a standard normal-form game against agent 1 in any cases. This would be natural when no knowledge is transferred. But once he is notified that there is some knowledge of which he was previously not aware, he may consider the reason why his opponent has told that to him. That is, he may consider, just as our (the modeler's) point of view, her telling the knowledge is as a result of her rational choice. If this is the case, his perception about the game structure is no longer just a normal-form game. He also has to take into account agent 1's reasoning about how he would have behaved if she did not tell the knowledge, for instance. Thus he would notice that he is now at some information set in an extensive-form game with unawareness. In the case of the Bach-Stravinsky-Mozart game, he might be able to consider agent 1's reasoning in the case of G^1 and G^2. In this case, this does not affect the result mentioned above because once he is notified about the Mozart concert, choosing it is his dominant strategy. But, in general, such an extension of the scope of agent 2's inference can lead to change in his rationalizable behavior due to the nature of extensive-form rationalizability: it deals with an agent's forward induction. Also, if agent 1 knows this, then her rationalizable strategies can also be different. But, even in such cases, agent 2 cannot infer anything that contains some knowledge of which he is still unaware. Thus, the extended scope depends on what knowledge is transferred at first. Taking into account these things generally in KT games would be a challenging task if we seriously consider agent 2's scope of inference, though it may result in a very complex form of extensive-form game with unawareness.

Other Types of Asymmetric Awareness and Knowledge Transfer under Them

It is the basic assumption of KT games above that agent 1 correctly perceives the objective game while agent 2 is unaware of something and his unawareness is a part of agent 1's view. This is only a particular class of asymmetric awareness in two-person games. In general, the framework of games with unawareness allows us to deal with infinite hierarchies of perceptions. Thus, for example we can describe such a situation that, in the Bach-Stravinsky-Mozart game, agent 2 actually perceives the objective game correctly and furthermore knows agent 1's belief about his unawareness, that is, he considers that the true game is G^1, she also views G^1 and she believes he believes G^2 is common knowledge. Since he now believes he has some knowledge that he believes she does not have, he may also be motivated to tell something to her, e.g., "I know there is the Mozart concert." Unlike the game of Figure 3 discussed above, this message of agent 2 must be credible for agent 1. (See Feinberg (2008) in which he studies this issue.) In this case, the situation becomes much more complicated because the both agents may be willing to tell something to each other. Therefore, in order to deal with other types of asymmetric awareness and knowledge transfer under them, it becomes necessary to identify what kind of communication can take place prior to playing the game and its influence to the agents' perceptions, which also would be a non-trivial task.

Furthermore, an agent can be aware of unawareness of herself. That is, she may notice that she might be unaware of something which she cannot identify. Games with unawareness have been extended to incorporate such awareness of unawareness (Heifetz et al., 2013; Halpern and Rêgo, 2014). Now consider the following situation. Agent 1 believes it is common knowledge that there is the Mozart concert, that is, G^1 in Figure 2 is common knowledge, while agent 2 believes G^2 is common knowledge. In this case, agent 1 would not be motivated to tell her opponent the fact that there is the Mozart concert since she believes he already knows it. Now, if agent 2 considers that G^2 might not be common knowledge and his opponent may know something of which he is unaware, what would he do? For instance, he might ask agent 1, "To my knowledge, there are concerts of Bach and Stravinsky, but do you know any other?" Then this question would be a surprise to her because she has considered he knows Mozart. But, anyway, she would be willing to tell him about the Mozart concert because the knowledge transfer can possibly lead to the Pareto-optimal outcome, where they both choose Mozart. Modeling this type of communication based on awareness of unawareness may result in a more complicated task, though it would be a challenging future work. (In systems science, the notion of systems intelligence refers to this kind of systemic feedback loops initiated by an agent having awareness of unawareness. See e.g. Hämäläinen and Saarinen (2006); Sasaki et al. (2014).)

APPENDIX

Our Definition of Rationalizability and Its Relation to Conventional Studies

We show that our definition of (extensive-form) rationalizability is equivalent to that defined originally by Heifetz et al. (2013). (We do not show the detail of their definition here because it is too lengthy.) In their model, each agent's strategy is defined as a combination of her actions at every information set she can perceive along some path in the game. Therefore, in a KT game, agent 1's strategy set is $\bar{A}_1 \times \prod_{k \in \{k' | G^{k'} \in \Gamma^1\}} A_1^k \times \prod_{l \in \{l' | G^{l'} \in \Gamma^2\}} A_1^l$ while that of agent 2 is $\prod_{l \in \{l' | G^{l'} \in \Gamma^2\}} A_2^l$. Each agent's rationalizable strategy is defined on the strategy set. It can be easily shown that the set of agent 2's rationalizable strategies is $\prod_{G^l \in \Gamma^2} R_2(G^l)$ since the process of iterated elimination of never-best-response strategies implies iterated elimination of never-best-response actions in each game in Γ^2, which means calculating (normal-form) rationalizability in these games. Thus it is equivalent to our setting that he just chooses his rationalizable action in every $G^l \in \Gamma^2$.

On the other hand, as for agent 1's rationalizable strategy, the following two statements are equivalent for $s_1^* = (x^*, a_1^*) \in S_1$: (i) s_1^* is her rationalizable strategy in our sense (i.e. $s_1^* \in R_1$); (ii) any strategies of agent 1 in $\{x^*\} \times \{a_1^*\} \times \prod_{k \in \{k' | G^{k'} \in \Gamma^1 \wedge G^{k'} \neq h(x)\}} A_1^k \times \prod_{l \in \{l' | G^{l'} \in \Gamma^2\}} A_1^l$, which we denote $\bar{R}_1(s_1^*)$, are rationalizable for her in the conventional sense. That is, in the conventional definition of rationalizability, she can choose any action at each information set other than h_0 and the normal-form game led by the first choice. The equivalence can be shown by the following argument. Suppose $s_1^* \in R_1$. This means, by definition, there exists$\sigma_2 \in \Sigma_2$ such that $Eu_1(s_1^*, \sigma_2) \geq Eu_1(s_1, \sigma_2)$ for any $s_1 \in S_1$. Thus, when she takes any strategy in $\bar{R}_1(s_1^*)$, she is sequentially rational at every information set under such a belief σ_2. Then such a strategy can survive the process of iterated elimination because agent 2's strategies cannot be eliminated any more as σ_2 only assigns probabilities to his actions included in his rationalizable strategy. Therefore it is her rationalizable strategy in the conventional sense. The converse can also be shown similarly. Hence our definition of rationalizability in KT games surely has the same spirit of it original definition.

Proof of Proposition 2

In a KT game, suppose $x^* = [\phi]$ for any $(x^*, a_1^*) \in R_1$ and the knowledge transfer cost is c. When the cost increases to $c'(> c)$, denote agent 1's expected utility function by $\overline{Eu_1^l}$ in any $G^l \in \Gamma^2$. In G^2, $\overline{Eu_1^2}(a_1, \delta_2) = Eu_1^2(a_1, \delta_2)$ for any $a_1 \in A_1^2$ and $\delta_2 \in \Delta A_2^2$, while, in any $G^l \in \Gamma^2 \setminus \{G^2\}$, $\overline{Eu_1^l}(a_1, \delta_2) = Eu_1^l(a_1, \delta_2) - (c' - c)$ for any $a_1 \in A_1^l$ and $\delta_2 \in \Delta A_2^l$. Therefore, in any $G^l \in \Gamma^2$, the set of agent 2's (normal-form) rationalizable actions does not change even when the cost increase occurs. Then, denote also her expected utility function in the extensive-form game which consists of h_0 and Γ^1 by $\overline{Eu_1}$ under the increased cost c'. Whenever she chooses $x \in \bar{A}_1$ other than $[\phi]$ at h_0, $\overline{Eu_1}(s_1, \sigma_2) = Eu_1(s_1, \sigma_2) - (c' - c)$ for any $\sigma_2 \in \Sigma_2$. (Note that Σ_2 does not change by the cost increase since agent 2's rationalizable actions does not change.) Thus, $x^* = [\phi]$ for any $(x^*, a_1^*) \in R_1$ under the increased cost as well. Hence the proposition holds.

Proof of Proposition 3

In a KT game, suppose there exists $x^* \neq [\phi]$ such that $(x^*, a_1^*) \in R_1$ for some $a_1^* \in A_1^k$, where $h(x^*) = G^k$, and the knowledge transfer cost is c. In a similar way as in the proof of Proposition 2, it can be shown that, in any $G^l \in \Gamma^2$, the set of agent 2's (normal-form) rationalizable actions does not change when the knowledge transfer cost decreases to $c'(< c)$. Then, whenever agent 1 chooses $x \in \bar{A}_1$ other than $[\phi]$ at h_0, $\overline{Eu_1}(s_1, \sigma_2) = Eu_1(s_1, \sigma_2) - (c' - c)$ for any $\sigma_2 \in \Sigma_2$, where $\overline{Eu_1}$ is defined in the same way as above. Thus, $(x^*, a_1^*) \in R_1$ under the decreased cost as well. Hence the proposition holds.

Proof of Proposition 4

In a KT game, for $x^* \in \bar{A}_1$ such that $h(x^*) = G^k$, suppose there exists $a_1^* \in A_1^k$ such that $s_1^* = (x^*, a_1^*) \in R_1$. By definition, this means that there exists $\sigma_2 \in \Sigma_2$ such that $Eu_1(s_1^*, \sigma_2) \geq Eu_1(s_1, \sigma_2)$ for any $s_1 \in S_1$. Let us denote, for any $G^l \in \Gamma^1$, $\max_{\delta_2 \in \Delta R_2(f(G^l))} (\max_{a_1 \in A_1^l} Eu_1^l(a_1, \delta_2))$ and $\min_{\delta_2 \in \Delta R_2(f(G^l))} (\max_{a_1 \in A_1^l} Eu_1^l(a_1, \delta_2))$ by $\max(G^l)$ and $\min(G^l)$, respectively. Now assume that there exists $G^{k'} \in \Gamma^1 \setminus \{G^k\}$ such that $\max(G^k) < \min(G^{k'})$. Let $h(x') = G^{k'}$. Then, for any $\sigma_2 \in \Sigma_2$, $Eu_1(s_1^*, \sigma_2) \leq \max(G^k)$ and there exists $a_1' \in A_1^{k'}$ such that $\min(G^{k'}) \leq Eu_1((x', a_1'), \sigma_2)$. Thus, $Eu_1(s_1^*, \sigma_2) < Eu_1((x', a_1'), \sigma_2)$. But this contradicts the definition above. Hence $\max(G^k) \geq \min(G^{k'})$ for any $G^{k'} \in \Gamma^1 \setminus \{G^k\}$.

Conversely, suppose that, for any $G^{k'} \in \Gamma^1 \setminus \{G^k\}$, $\max(G^k) \geq \min(G^{k'})$. Denote, for any $G^l \in \Gamma^1$, the argument of $\max(G^l)$ by $\mathrm{argmax}(G^l)$. Likewise, $\mathrm{argmin}(G^l)$ is defined. Now consider a belief $\sigma_2 \in \mathrm{argmax}(G^k) \times \prod_{G^{k'} \in \Gamma^1 \setminus \{G^k\}} \mathrm{argmin}(G^{k'})\ (\subseteq \Sigma_2)$ and agent 1's strategy $s_1^* = (x^*, a_1^*) \in S_1$, where $a_1^* \in \mathrm{argmax}_{a_1 \in A_1^k} Eu_1^k(a_1, \delta_2)$ in which the non-zero probabilities assigned by $\delta_2 \in \Delta A_2^k$ is equal to the probabilities on $R_2(f(G^k))$ specified in σ_2. Then $Eu_1(s_1^*, \sigma_2) \geq Eu_1((x^*, a_1), \sigma_2)$ for any $a_1 \in A_1^k$ due to the definition of a_1^*. Also $Eu_1(s_1^*, \sigma_2) \geq Eu_1((x, a_1), \sigma_2)$ for any $x \in \bar{A}_1 \setminus \{x^*\}$ and $a_1 \in A_1^{k'}$ when $h(x) = G^{k'}$ because the right-hand side is equal to $\max(G^k)$ and the left-hand side is not greater than $\min(G^{k'})$ when $h(x) = G^{k'}$. To sum up, $Eu_1(s_1^*, \sigma_2) \geq Eu_1(s_1, \sigma_2)$ for any $s_1 \in S_1$, which means $s_1^* \in R_1$. Hence the proposition holds.

Proof of Proposition 5

In a KT game, $\max(G^l) \geq \min(G^l)$ in any $G^l \in \Gamma^1$, where $\max(G^l)$ and $\min(G^l)$ are defined in the same way as in the proof of Proposition 4. Thus, for $x^* \in \bar{A}_1$ such that $h(x^*) = G^k$, if $\min(G^k) \geq \min(G^{k'})$ for any $G^{k'} \in \Gamma^1 \setminus \{G^k\}$, then $\max(G^k) \geq \min(G^{k'})$ for any $G^{k'} \in \Gamma^1 \setminus \{G^k\}$. Due to Proposition 4, this implies there exists $a_1^* \in A_1^k$ such that $(x^*, a_1^*) \in R_1$. Hence the proposition holds.

ACKNOWLEDGEMENT

This research is supported by JSPS KAKENHI Grant-in-Aid for Young Scientists (B) 15K16292.

REFERENCES

Bandyopadhyay, S. and Pathak, P. (2007). Knowledge sharing and cooperation in outsourcing projects - A game theoretic analysis, *Decision Support Systems*, 43(2):349-358.

Battigalli, P. (1997). On rationalizability in extensive games, *Journal of Economic Theory*, 74(1):40-61.

Bennett, P. G. (1977). Toward a theory of hypergames, *Omega*, 5:749-751.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission, *Econometrica*, 50(6):1431-1451.

Feinberg, Y. (2008). Meaningful talk, in *New perspectives on games and interaction*, (Krzysztof, R. A. and van Rooij, R., ed.), Amsterdam University Press, Amsterdam.

Halpern, J. Y. and Rêgo, L. C. (2014). Extensive games with possibly unaware players, *Mathematical Social Sciences*, 70:42-58.

Hämäläinen, R. P. (1981). On the cheating problem in Stackelberg games, *International Journal of Systems Science*, 12(6):753-770.

Hämäläinen, R. P. and Saarinen, E. (2006). Systems intelligence: A key competence in human action and organization life, *The SoL Journal*, 7(4):17-28.

Harsanyi, J. C. (1967). Games with incomplete information played by Bayesian players, *Management Science*, 14: 159-182, 320-334, 486-502.

Heifetz, A., Meier , M. and Schipper, B. C. (2013). Dynamic unawareness and rationalizable behavior, *Games and Economic Behavior*, 81:50-68.

Li, Y. M. and Jhang-Li, J. H. (2010). Knowledge sharing in communities of practice: A game theoretic analysis, *European Journal of Operational Research*, 207(2):1052-1064.

Myerson, R. B. (1991). *Game Theory*, Harvard University Press, Cambridge.

Nonaka, I. and Takeuchi, H. (1995). *The Knowledge-creating Company*, Oxford University Press, New York.

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

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection, *Econometrica*, 52:1029-1050.

Perea, A. (2012). *Epistemic Game Theory: Reasoning and Choice*, Cambridge University Press, Cambridge.

Samaddar, S. and Kadiyala, S. S. (2006). An analysis of interorganizational resource sharing decisions in collaborative knowledge creation, *European Journal of Operational Research*, 170(1):192-210.

Sasaki, Y. and Kijima, K. (2012). Hypergames and Bayesian games: A theoretical comparison of the models of games with incomplete information, *Journal of Systems Science and Complexity*, 25(4):720-735.

Sasaki, Y., Hämäläinen, R. P. and Saarinen, E. (2014). Modeling systems of holding back as hypergames and their connections with systems intelligence, *Systems Research and Behavioral Science*, doi: 10.1002/sres.2276.

Schipper, B. C. (2014). Unawareness - A gentle introduction to both the literature and the special issue, *Mathematical Social Sciences*, 70:1-9.

Spence, M. (1973). Job market signaling, *Quarterly Journal of Economics*, 87(3):355-374.

Wang, M., Hipel, K. W. and Fraser, N. M. (1988). Modeling misperceptions in games, *Behavioral Science*, 33:207-223.