Iterative Information Update and Stability of Strategies

Takuya Masuzawa[1] and Koji Hasebe[2]

[1] Faculty of Economics, Osaka University of Economics
2-2-8 Osumi, Higashiyodogawa-ku, Osaka, 533-8533 Japan
masuzawa@osaka-ue.ac.jp

[2] Graduate School of Systems and Information Engineering, University of Tsukuba
1-1-1 Tennoudai, Tsukuba-shi, 305-8573 Japan
hasebe@iit.tsukuba.ac.jp

Abstract. In this paper, we investigate processes of iterative information update due to Benthem (*International Game Theory Review*, vol.9, pp.13-45, 2007), who characterized existent game-theoretic solution concepts by such processes in the framework of Plaza's public announcement logic. We refine this approach and make clearer the relationship between stable strategies and information update processes. We first extend Plaza's logic then demonstrate the conditions under which a stable outcome is determined independently of the order of iterative information update. This result gives an epistemic foundation for the order independence of iterated elimination of disadvantageous strategies.
Keywords: game theory, epistemic logic, information update, public announcement, stable strategy

1 Introduction

Logical analysis of game-theoretic solution concepts is an active trend in studies on dynamic epistemic logic [8] and its variants (cf. also [4]). In that domain, researchers consider processes of updating each agent's knowledge, which changes the epistemic models for a given game. In particular, by means of Plaza's public announcement logic [15], Benthem [2] analyzed game theoretic solution concepts by considering an information updating process. In his setting, based on a decision criterion every player first chooses a strategy as a tentative decision, and then each player's information about the situation is iteratively updated depending on the other players' tentative decisions. The information update may bring about a failure of the criterion to keep the tentative decision. That is, the information update may suggest that another choice is preferable to some players, and they may change their tentative decisions. In this setting, Benthem considered a stability of strategy, which we call *iterative updatability* in this paper. Roughly speaking, a tentative decision is iteratively updatable if no player changes it for any number of public announcements that no player has changed his tentative decision so far. Through this notion, Benthem characterized existent game-theoretic solution concepts.

The purpose of our paper is to refine the approach of [2] and make clearer the relationship between stable strategies and information update processes. For this purpose, we also introduce a logic based on Plaza's public announcement logic [15]. In order to describe game-theoretic components, such as players' evaluation of a situation and intensions for their choices, we first extend Plaza's logic so that all the atomic formulas are classified into three categories: *information-invariant* (or called invariant), *information-monotonic* (or called monotonic), and the others. An invariant atomic formula represents a statement whose truth value does not change after any information update. On the other hand, a monotonic atomic formula represents a statement whose truth persists after any information update if it is previously true. In this sense, in Plaza's logic all the atomic formulas are regarded as invariant. This extension provides an enough expressive power to describe the game-theoretic components whose truth might change after some announcement. Then we also introduce an extended Kripke-style possible world semantics and show the soundness and completeness for this syntax.

Next we formalize the notion of iterative updatability in terms of our logic to investigate its properties, especially in relation with information-monotonicity. We show that if announced statements are information monotonic, then the order of announced statements does not affect the information updatability. This result explains the order independence of iterated elimination of disadvantageous strategies. Moreover, we show that the iterative process of information update preserves logical implication between two statements if one of them is monotonic. This theorem is useful when we compare different information updating processes.

Finally, we also demonstrate how to apply our results to game-theoretic situations. The first example is an exchange economy with asymmetric information. By iterative updatability, we explain how the quality of a good, which is initially private information, is revealed to other market participants. As a second application, following [2], we reexamine iterated elimination of disadvantageous strategies. By means of our refined framework, we demonstrate properties such as order independence, comparison of two criteria, and its relation to Nash equilibria.

Paper organization. Section 2 introduces the syntax and semantics of our logic. Section 3 formulates the notion of iterative updatability, and then shows its properties. Section 4 demonstrates how to analyze a game-theoretic situation by iterative updatability. Finally, Section 5 concludes the paper.

2 Public announcement logic

In this section we introduce our inference system, which is an extension of Plaza's public announcement logic [15].

2.1 Language

Let N be a set of players and P be an infinitely countable set of atomic formulas (denoted by the symbols p, q, $\ldots$). Here, we introduce the classification of atomic formulas into the following three types: *information-invariant* formulas (denoted by Q^*), *information-monotonic* formulas (denoted by Q), and the others, where $Q^* \subseteq Q \subseteq P$. This classification is the difference between our logic and that of Plaza's. Throughout the paper, for brevity, Q^* and Q are often called *invariant* and *monotonic* atomic formulas, respectively.

Formulas (denoted by φ, $\psi, \ldots$) are constructed by the following grammar, which is the same as in Plaza's logic.

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \Rightarrow \varphi \mid \varphi \leftrightarrow \varphi \mid K_i\varphi \mid \langle\varphi\rangle\varphi$$

Intuitively, a formula of the form $K_i\varphi$ means that player i knows φ, while a formula of the form $\langle\varphi\rangle\psi$ means that ψ holds after the true statement φ has been publicly announced.

2.2 Axiomatic system

The axioms and inference rules of our system are as follows.

Axioms for the propositional tautology and epistemic operators:
A1 Every axiom of the propositional tautology is an axiom.
A2 $K_i(\varphi \Rightarrow \psi) \Rightarrow (K_i\varphi \Rightarrow K_i\psi)$.
A3 $K_i\varphi \Rightarrow \varphi$.

Axioms for the public announcement operator:
P1 $\langle\varphi\rangle\psi \Rightarrow \varphi$.
P2 $\langle\varphi\rangle\neg\psi \leftrightarrow \varphi \wedge \neg\langle\varphi\rangle\psi$.
P3 $\langle\varphi\rangle(\psi \wedge \chi) \leftrightarrow \langle\varphi\rangle\psi \wedge \langle\varphi\rangle\chi$.
P4 $\langle\varphi\rangle K_i\psi \leftrightarrow \varphi \wedge K_i(\varphi \Rightarrow \langle\varphi\rangle\psi)$.
P5 $\langle\varphi\rangle\langle\psi\rangle\chi \leftrightarrow \langle\langle\varphi\rangle\psi\rangle\chi$.

A theorem of our system is inductively defined as follows.

R1 If φ is an axiom, then $\vdash \varphi$.
R2 Modus ponens: If $\vdash \varphi$ and $\vdash (\varphi \Rightarrow \psi)$, then $\vdash \psi$.
R3 Necessitation: If $\vdash \varphi$, then $\vdash K_i\varphi$.
R4 Substitution of equals for public announcement:
If $\vdash \varphi \leftrightarrow \psi$, then $\vdash \langle\varphi\rangle\chi \leftrightarrow \langle\psi\rangle\chi$ and $\vdash \langle\chi\rangle\varphi \leftrightarrow \langle\chi\rangle\psi$.
R5 Invariance for null information: If $\vdash \varphi$, then $\vdash \psi \leftrightarrow \langle\varphi\rangle\psi$.

Note that the system composed of A1-A3 and R1-R3 is the usual multimodal propositional logic. On the other hand, axioms P1-P5 and inference rules R4 and R5 are used for reasoning about public announcements.

In addition to all the axioms and inference rules introduced above, Plaza's logic includes the axiom that $\langle\varphi\rangle q \leftrightarrow \varphi \wedge q$ for any atomic formula q. Instead of this axiom, we introduce the following axiom P0 and inference rule R6 to formalize the notions of invariance and monotonicity, respectively.

Axiom for invariance:
P0 $\langle\varphi\rangle q \leftrightarrow \varphi \wedge q$ for all $q \in Q^*$.

Inference rule for monotonicity:
R6 If $\vdash \varphi \Rightarrow \psi$, then $\vdash \langle\varphi\rangle\neg q \Rightarrow \langle\psi\rangle\neg q$ for all $q \in Q$.

Axiom P0 means that the truth value of any invariant atomic formula does not change after any announcement. On the other hand, inference rule R6 means that the truth of any monotonic atomic formula is preserved after any announcement if it is initially true.

Here, we make some remarks on this extension of Plaza's logic. As mentioned in Section 1, in the framework of Plaza's logic, the truth value of an atomic formula does not change after any information update. In other words, Plaza's logic is a special case in which $Q^* = Q = P$ (i.e., any atomic formula is treated as invariant). To describe our target situation, however, in which the preferences or intentions of players may change depending on information updates, we require some distinction between invariant and non-invariant statements. Our logic is a minimal extension to solve this problem.

From the syntactic point of view, Plaza's formalism allows us to translate any formula into an equivalent one without the public announcement operator. Plaza used this property to prove some logical meta-theorems, such as the completeness theorem [15]. On the other hand, this property does not hold for our logic, but through semantic extension we can also prove its completeness, which is shown in the next subsection.

Regarding the notion of monotonicity, we can extend it to general formulas. Let $\mathcal{M}$ be a set of formulas, each of which (say, χ) satisfies the following condition: if $\vdash \varphi \Rightarrow \psi$, then $\vdash \langle\varphi\rangle\neg\chi \Rightarrow \langle\psi\rangle\neg\chi$. Clearly, this condition is a natural extension of R6. For $\mathcal{M}$, the following proposition holds.

Proposition 1. *$\mathcal{M}$ is closed under the following operations:*

1. *If $\chi_1 \in \mathcal{M}$ and $\chi_2 \in \mathcal{M}$, then $(\chi_1 \wedge \chi_2)$, $(\chi_1 \vee \chi_2)$ are also in $\mathcal{M}$.*
2. *If $\chi_1 \in \mathcal{M}$, then $K_i\chi_1$ is also in $\mathcal{M}$.*

These operations do not fully characterize $\mathcal{M}$. For example, we can show that if $\chi_1 \in \mathcal{M}$ and $\chi_2 \in \mathcal{M}$, then $\neg\chi_1 \Rightarrow \langle\neg\chi_1\rangle\chi_2$ are also in $\mathcal{M}$. Further, note that all theorems are monotonic.

2.3 Semantics

Our language can be interpreted in standard models for epistemic logic, except for interpretation of invariance and monotonicity.

A Kripke-model M is a triple $(W, (R)_{i\in N}, v)$, where W is a set of states, R_i an accessibility relation over W for player $i \in N$, and $v : P \times W \times 2^P \to \{1, 0\}$ is an assignment function. We assume that R_i is reflexive. We impose the following conditions on v:

Invariance of Q^* : For any $q \in Q^*$, $v(q^*, w, X') = v(q^*, w, X'')$ for any $w \in W$, and for any $X', X'' \subseteq W$.

Monotonicity of Q : If $X' \subseteq X'' \subseteq W$ and $v(q, w, X'') = 1$, then $v(q, w, X') = 1$ for any $w \in W$.

Here, remember that Q^* and Q are the sets of information-independent and information-monotonic atomic formulas, respectively.

The difference between usual Kripke semantics and ours lies in the interpretation of atomic formulas. In usual Kripke semantics, the truth value of an atomic formula is determined for each state. By means of this assignment function, however, we cannot treat a situation in which the truth value of an atomic formula in a certain state may change after some public announcement. In order to formalize such dynamism, our idea is to extend the assignment function so that the truth value of an atomic formula in a state is determined by a set of states $X \subseteq W$. Then, we define the truth values of atomic formulas for such sets $X, X', X'', \ldots \subseteq W$, which are obtained after public announcements.

Definition 1 (Truth conditions). *The truth value of a formula in state* w *of* $M = (W, (R)_{i \in N}, v)$ *is defined as follows. Here,* $(w, M) \models \varphi$ *denotes that* φ *is true in state* w *of model* M.

1. *For an atomic formula* $q \in P$, $(w, M) \models q$ *iff* $v(q, w, W) = 1$.
2. *For all* ψ *and* φ, *each of* $\neg\psi$, $\psi \Rightarrow \varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, *and* $\varphi \leftrightarrow \psi$ *is true in* w *of* M *iff it is true in the truth table of* ψ *and* φ.
3. *For all* ψ *and* $i \in N$, $(w, M) \models K_i\psi$ *iff for all* w' *such that* wR_iw', $(w', M) \models \psi$.
4. *For all* ψ *and* φ, $(w, M) \models \langle\psi\rangle\varphi$ *iff* $(w, M) \models \psi$ *and* $(w, M|\psi) \models \varphi$, *where* $M|\psi = (W|\psi, (R_i|\psi)_{i \in N}, v|\psi)$, $W|\varphi = \{w' \in W : (w, M) \models \psi\}$, *and* $R_i|\psi$ *and* $v|\psi$ *are the restrictions of* R_i *and* v *to* $W|\psi$, *respectively.*

We also define that φ is valid (denoted by $\models \varphi$) iff $(w, M) \models \varphi$ for any model M and for any state w.

For the syntax and semantics introduced so far, we can prove the following soundness and completeness theorem by similar argument as in [15].

Theorem 1 (Soundness and Completeness). *For any formula* φ, $\vdash \varphi$ *iff* $\models \varphi$.

Proof. Due to space limitations, we only show the completeness. It suffices to show that for all φ, if $\nvdash \neg\varphi$ then φ is satisfiable. Let $S(\varphi)$ be the set of all subformulas of φ and we say that a set U of formulas is $S(\varphi)$-*maximal consistent* if (1) for all $\varphi' \in S(\varphi)$ either one of φ' and $\neg\varphi'$ is in U, (2) for all $\varphi_1 \in U$ there exists $\varphi_2 \in S(\varphi)$ such that either $\varphi_1 = \varphi_2$ or $\varphi_1 = \neg\varphi_2$, (3) if $\varphi_1, \varphi_2, \ldots, \varphi_m \in U$ then $\nvdash \neg(\wedge_i^m \varphi_i)$. Let $M = (W, (R_i), v)$ be a Kripke-model such that

- $W = \{U : U$ *is a* $S(\varphi)$*-maximal consistent set.*$\}$,
- UR_iU' iff φ is in U' for all φ such that $K_i\varphi \in U$,
- for all $p \in P$, $v(p, U, X) = 1$ iff $p \in U$,

– for all $Y \neq X$ and all $p \in (P \setminus Q)$, $v(p, U, Y) = 1$ iff $Y = \{U' \in W : \varphi_1 \in U'\}$, $\langle \varphi_1 \rangle p \in U$ for some $\langle \varphi_1 \rangle p \in S(\varphi)$,
– for all $Y \neq X$ and all $q \in Q$, $v(q, U, Y) = 1$ iff either $q \in U$ or $Y \subset \{U' \in X | \varphi_1 \in U'\}$, $\langle \varphi_1 \rangle q \in U$ for some $\langle \varphi_1 \rangle q \in S(\varphi)$.

For this model, by induction on φ we can show that $\varphi \in U$ iff $(U, W) \models \varphi$ for all φ. □

As an application, we have the following result:

Proposition 2. *If $Q = \emptyset$, for any formula $\varphi \in \mathcal{M}$, either $\vdash \varphi$ or $\vdash \neg\varphi$.*

Proof. Suppose to the contrary that there exists $\varphi \in \mathcal{M}$ such $\nvdash \varphi$, $\nvdash \neg\varphi$.

First, we consider the case where for any Kripke-model whose set of the states is singleton, φ is false. By the completeness theorem, we have a finite model $M_1 = (W_1, (R_{i,1}), v_1)$ and $w \in W_1$ such that for some state w, $(w, M_1) \models \varphi$. We can assume $|W_1|$ is minimal. Then, we can find a literal p such that (1) $(w, M_1) \models p$ and (2) $|W_1| > |W_1|p|$. Thus, $(w, M_1) \models \varphi \wedge \langle p \rangle \neg\varphi$, which contradicts the monotonicity of φ.

Now, consider the second case where there exists a Kripke-model M_2 such that the set of states is singleton and that φ is true. By the completeness theorem, we have a finite model $M_3 = (W_3, (R_{i,3}), v_3)$ and $w \in W_3$ such that $(w, M_3) \models \neg\varphi$. Without any loss of generalities, assume that the number of state is minimal.

If $v(q, w, W_3) \neq v(q, w', W_3)$ for some q and $w' \neq w$ then $W_3|q \neq W_3$ and we have, from the minimality, $w, M_3|q \models \varphi$, which contradicts the monotonicity of φ. Then we can assume that $v(q, w, W_3) = v(q, w', W_3)$ for all q and $w' \neq w$. Then in any state of W_3, the truth-value of any formula is equivalent. From the minimality, $|W_3| = 1$. Let us construct a model M_4 as follows: $W_4 = \{a_1, a_2, \ldots, a_n\}$; $xR_{i,4}y$ iff $x = y$. Let X_k denote $\{a_1, a_2, \ldots, a_k\}$. If $l < k$ then $v_3(q, a_l, X_k)$ is determined as M_2; otherwise it is determined as M_3. In this model, inductively we can prove that for all $\deg(\psi) \leq k - l$, $(a_l, X_k) \models \psi$ if $(w, M_2) \models \psi$. (Here, deg stands for the number of logical connectives.) Thus, in the case of $n = \deg(\varphi) + 1$, $(a_1, X_n) \models \varphi$. On the other hand, there exists a literal formula q such that it is true in M_2 but is false in M_3. Let us define $q_k = \langle q_{k-1} \rangle q$. Then, $X_n|q_{n-1} = X_1$. Thus, $a_1, M_4|q_{n-1} \models \neg\varphi$. □

3 Iterative updatability

By the logic introduced in the previous section, we define the iterative updatability and present useful properties as well as an economic example.

3.1 Definition

We consider the condition under which players maintain the given tentative decision. Note that even if the player's criterion to maintain the tentative decision is satisfied at first, the announcement of this fact may bring about a failure of the criterion and the change of the tentative decision. Thus, additional conditions are required for the tentative decision to be maintained subsequently.

To see that, consider a situation with two-players where each of them chooses a tentative decision. Let φ_i denote a condition for player i (for $i \in \{1, 2\}$) to maintain his tentative decision. The tentative decision may be changed even if φ_i is true for all i at first. Suppose that a player, i, observes that the other player, j (for $j \neq i$), does not change his tentative decision, namely φ_j. Then, even if each player does not change his tentative decision at first, the observation can change his information and turn φ_i into a false statement. In terms of our logic, $\langle\varphi_j\rangle\varphi_i$ may be false while φ_j and φ_i are both true. Further, in turn, observing that i does not change the tentative decision regardless of his observation on j, j might change his tentative decision. That is, $\langle\varphi_j\rangle\langle\varphi_i\rangle\varphi_j$ might be false even if $\langle\varphi_j\rangle\varphi_i$. Similarly, for the tentative decision to be maintained, we also require $\langle\varphi_j\rangle\langle\varphi_i\rangle\langle\varphi_j\rangle\varphi_i$, $\langle\varphi_j\rangle\langle\varphi_i\rangle\langle\varphi_j\rangle\langle\varphi_i\rangle\varphi_j$, and so on. In summary, for a tentative decision to be unchanged, it is necessary to take such an iterated information update process into consideration. More formally, a tentative decision is stable only if $\langle\varphi_i\rangle\varphi_j$, $\langle\varphi_j\rangle\langle\varphi_i\rangle\varphi_j$, $\langle\varphi_i\rangle\langle\varphi_j\rangle\varphi_i$ and similar propositions are all true.

Let us formalize the robustness against information updates, *iterative updatability.* We mean by $\mathcal{C}$ a set of formulas such that each formula, φ in $\mathcal{C}$, represents a given sufficient condition for a player to change his tentative decision. Let $\mathcal{C}^-$ be a set of negations of $\mathcal{C}$: $\mathcal{C}^- = \{\neg\varphi : \varphi \in \mathcal{C}\}$. Then we can formulate our target condition by $\mathcal{C}^-$:

Definition 2 (Iterative Updatability). *We say that a set of formulas $\mathcal{C}^-$ is iteratively updatable in a state of a Kripke model (w, M) if, for all finite sequences of formulas, $\psi_1, \psi_2, \ldots, \psi_k \in \mathcal{C}^-$,*

$$(w, M) \models \langle\psi_1\rangle\langle\psi_2\rangle \ldots \langle\psi_{k-1}\rangle\psi_k.$$

Note that Benthem [2] originally focused on the case when $|\mathcal{C}| = 1$.

If $\mathcal{C}^-$ is iteratively updatable, then each statement is true and any iterated announcement maintains the statements' truth values as true. Then, the condition to discuss is the *iterative updatability* of $\mathcal{C}^-$. We shall discuss properties of the iterative updatability in the following subsection.

3.2 Basic properties

In this subsection, we show that if $\mathcal{C}$ consists of information-monotonic statements, the iterative updatability of $\mathcal{C}^-$ has useful properties.

First property is related to the order of announced statements. For iterative updatability, it does not suffice to consider only the case when specific statements are announced in fixed turns, even if they are iteratively announced. On the other hand, in the case of $\mathcal{C} \subseteq \mathcal{M}$, we can prove the following theorem.

Theorem 2 (Single Sequence Property). *Suppose that $\mathcal{C}$ is a subset of $\mathcal{M}$. Then, $\mathcal{C}^-$ is iteratively updatable if and only if there exists a sequence $\psi_1, \psi_2, \ldots$ in $\mathcal{C}^-$ such that (1) every $\varphi \in \mathcal{C}^-$ appears in the sequence infinite times and (2) $\langle\psi_1\rangle\langle\psi_2\rangle \ldots \langle\psi_{k-1}\rangle\psi_k$ is true for all $k = 1, 2, \ldots$.*

Proof. Let $\varphi_1, \varphi_2, \ldots, \varphi_k$ be any finite sequence in $\mathcal{C}^-$. Define F_k by $F_1 = \psi_1$ and $F_k = \langle F_{k-1} \rangle \psi_k$. Let $\psi_{l1}, \psi_{l2}, \ldots, \psi_{lk}$ be a subsequence of $\psi_1, \psi_2, \ldots$ such that $\psi_{l1} = \varphi_1, \psi_{l2} = \varphi_2, \ldots, \psi_{lk} = \varphi_k$. From monotonicity of $\neg\varphi_1$, we have $\vdash F_{l1} \Rightarrow \varphi_1$. Thus, by P1, $\vdash F_{l2-1} \Rightarrow \varphi_1$. From the monotonicity, $\vdash \langle F_{l2-1} \rangle \psi_{l2} \Rightarrow \langle \varphi_1 \rangle \varphi_2$. Through repetition, we have $\vdash F_{lk} \Rightarrow \langle \varphi_1 \rangle \langle \varphi_2 \rangle \ldots \langle \varphi_{k-1} \rangle \varphi_k$. □

The second property we address is related to logical implication and iterative updatability. Consider the two sets of criteria for players to change their tentative decision, $\mathcal{C}$ and $\mathcal{C}'$. The iterative updatability of $\mathcal{C}$ does not imply that of $\mathcal{C}'$ even if $\mathcal{C}$ consists of stronger conditions than those of $\mathcal{C}'$. On the other hand, if they are information-monotonic then logical implication is preserved.

Theorem 3 (Comparison Theorem). *Let $\psi_1, \psi_2 \ldots$ and $\varphi_1, \varphi_2, \ldots$ be two sequences of formulas such that for all $i = 1, 2, \ldots, n$, $\vdash \psi_i \Rightarrow \varphi_i$. Assume that one of $\{\neg\varphi_i, \neg\psi_i\}$ is in $\mathcal{M}$ for all $i = 1, 2, \ldots, n$,. Then,*

$$\vdash \langle \psi_1 \rangle \langle \psi_2 \rangle \ldots \langle \psi_{n-1} \rangle \psi_n \Rightarrow \langle \varphi_1 \rangle \langle \varphi_2 \rangle \ldots \langle \varphi_{n-1} \rangle \varphi_n.$$

Proof. The theorem trivially holds for $n = 1$. Let F_n and G_n be inductively defined by $F_1 = \psi_1$, $G_1 = \varphi_1$, $F_k = \langle F_{k-1} \rangle \psi_k$, and $G_k = \langle G_{k-1} \rangle \varphi_k$ for $k > 1$. Suppose that $\vdash F_{k-1} \Rightarrow G_{k-1}$. Consider the case when $\neg\psi_k \in \mathcal{M}$. From the monotonicity, $\vdash \langle F_{k-1} \rangle \psi_k \Rightarrow \langle G_{k-1} \rangle \psi_k$. Since $\vdash \psi_k \Rightarrow \varphi_k$, we have $\vdash \langle G_{k-1} \rangle \psi_k \Rightarrow \langle G_{k-1} \rangle \varphi_k$. Then, $\vdash \langle F_{k-1} \rangle \psi_k \Rightarrow \langle G_{k-1} \rangle \varphi_k$. For the case of $\neg\varphi_k \in \mathcal{M}$, we can also prove it by a similar discussion. □

Since the monotonicity of only one of two formulas is assumed, this theorem is applicable to analyzing the case when the target condition is a non-monotonic statement. It might be unclear how the truth value of the target statement is changed by information updates, and thus, we sometimes focus on the sufficient or necessary condition, which is logically clearer than the target condition. If the necessary or sufficient condition is information-monotonic, then iterative updatability preserves both sufficiency and necessity. Note that all the theorems presented here still hold for any extended axiomatic system.

3.3 An exchange economy

To illustrate our concept and results above, we present an exchange market model with two buyers $B1, B2$, and one seller S. Each buyer has money, and the seller has one indivisible good, which might be of high quality or low quality. The seller's utility depends only on the money he has. When a buyer knows the quality of the good, he evaluates a high-quality good at \$8 and a low-quality good at \$4.[3] On the other hand, if he does not know the quality, then he evaluates the good at \$6.

[3] This preference is represented by the utility function $u(x_h, x_l, -p) = 4 \min\{2, 2x_h + x_l\} - p$, where x_h is the quantity of a high-quality good, x_l is that of a low-quality good, and p is the payment.

We consider a two-state model such that $W = \{h, l\}$ and R_i for $i \in \{S, B1, B2\}$, defined by

– for $B1$: $hR_{B1}h$, $hR_{B1}l$, $lR_{B1}l$, $lR_{B1}h$,
– for S and $B2$: hR_Sh, lR_Sl, $hR_{B2}h$, $lR_{B2}l$.

In state h, the good is of high quality, while in state l, it is of low quality. Then, $B2$ and S know the quality of the good, while $B1$ does not know it at first.

Let us consider the case when the players are going to do the following transactions denoted by

T0: S sells to $B1$; $B1$ pays \$5 to S.

On the other hand, we assume that alternative transactions are possible for each pair if it can be done by themselves. Especially, the pair of $B2$ and S can choose an alternative transaction, denoted by

T1: S sells to $B2$; $B2$ pays \$6 to S.

Further, each individual can decide not to participate in any transaction.

In each state, T0 is not acceptable to players in the sense that some player eventually deviates from T0. In state h, $B2$ and S can increase their payoffs from $(0, 5)$ to $(2, 6)$ by changing the tentative decision to T1, and thus, they should do so. On the other hand, in state l, no player attempts to change the tentative decision at first. By observing that either $B2$ or S has not changed his tentative decision, however, $B1$ deduces that the quality of the good is low. Then, his evaluation of the good is changed to \$4. Paying \$5 is no longer profitable for him, and thus, he attempts to interrupt T0 and does not participate in any other transaction.

Iterative updatability we introduced is a formalization of the acceptability in this situation. Let $\mathcal{C} = \{\neg q_i : i = B1, B2, s\}$, where q_i denotes that 'i maintains the tentative decision, namely, agreeing to T0'. Further, let φ be the statement that T1 increases the utilities of $B2$ and S from T0, and let ψ indicate that interrupting T0 increases the utility of $B1$. We can assume that $\vdash q_j \Rightarrow \neg K_j\varphi$ for $j = B2, S$, and that $\vdash q_{B1} \Rightarrow \neg K_{B1}\psi$. Then, in our semantics, $\mathcal{C}^-$ is not iteratively updatable. To see this, first note that $\langle \neg K_{B2}\varphi \rangle \neg K_{B1}\psi$ is false in any state. Further, $K_{B2}\varphi$ and $K_{B1}\psi$ are all monotonic, and thus, by the Comparison Theorem, $\langle q_{B2} \rangle q_{B1}$ is also false.

4 Iterated elimination of strategies

In this section, as an application of our results in the previous section, we examine iterated elimination (IE) of disadvantageous strategies in our framework.

4.1 Definition

Let X^i be a finite strategy set of $i \in N$. We abbreviate $\prod_{i \in S} X^i$ by X^S for all $S \subseteq N$. By x^S we denote a typical element of X^S for all $S \subseteq N$, and by

$(x^S, x^{N\setminus S})$ we denote an element in X^N such that the projections into X^S and $X^{N\setminus S}$ are x^S and $x^{N\setminus S}$, respectively. Further, let $u^i : X^N \to \Re$ be the utility function of $i \in N$.

We introduce three base criteria for elimination of disadvantageous strategies. A strategy $x^i \in X^i$ is *strictly dominated* iff $\exists y^i \in X^i\ \forall x^{N\setminus\{i\}} \in X^{N\setminus\{i\}}$: $u^i(y^i, x^{N\setminus\{i\}}) > u^i(x^i, x^{N\setminus\{i\}})$. It is *weakly dominated* iff $\exists y^i \in X^i\ \forall x^{N\setminus\{i\}} \in X^{N\setminus\{i\}}$: $u^i(y^i, x^{N\setminus\{i\}}) \geq u^i(x^i, x^{N\setminus\{i\}})$ with at least one strict inequality. Moreover, it is a *never best response* iff $\forall x^{N\setminus\{i\}} \in X^{N\setminus\{i\}}\ \exists y^i \in X^i : u^i(y^i, x^{N\setminus\{i\}}) > u^i(x^i, x^{N\setminus\{i\}})$. The survivors of iterated elimination (SIE), $\boldsymbol{X}^{*N}$, is defined by the following algorithm for any of the base criteria for elimination:

begin for all $i \in N$, $\boldsymbol{X}^{*i} := X^i$;
  while there exist $j \in N$ and $x^j \in X^{*j}$ that
  meets the base criterion for elimination in the reduced game,
  the strategy set of which is restricted to $\boldsymbol{X}^{*N}$
    begin for such j and x^j **do** $\boldsymbol{X}^{*j} := \boldsymbol{X}^{*j} \setminus \{x^j\}$ **end**
end

The SIE of strictly or weakly dominated strategies can be traced back to Gale, et al.[6], Gale [7], and Luce and Raiffa [13]. Bernheim [3] and Pearce [14] independently discussed IE of never best response strategies[4].

4.2 Epistemic characterization

These strategies have been characterized by static epistemic models. First, Pearce [14] characterized SIE of never best response as those chosen in a situation in which it is common knowledge that each player maximizes his utility with respect to his prior belief about other players' strategy choices as well as the structure of the game. Tan and Werlang [17] refined this conjecture and gave a formal proof in a Bayesian framework. Samuelson [16] considered a situation in which every player chooses a subset of strategies as admissible ones. He pointed out that it is impossible to characterize the set of SIE of weakly dominated strategies as those chosen in a situation in which it is common knowledge that each player's admissible strategies are weakly non-dominated strategies with respect to his belief.

Benthem [2] initiated an approach by public announcement logic to these concepts. Let $x^N \in X^N$ be a given status quo in which each player j chooses x^j as a tentative decision. For all $y^i \in X^i$, let $f(y^i)$ denote the statement 'for all $x^N \in X^N$, $u^i(y^i, x^{N\setminus\{i\}}) > u^j(x^i, x^{N\setminus\{i\}})$ if x^N is the tentative decision'. Further, we denote by $g(y^i)$ the statement 'for all $x^N \in X^N$, $u^i(y^i, x^{N\setminus\{i\}}) > u^j(x^i, x^{N\setminus\{i\}})$ if x^N is the tentative decision'. Here, we do not assume any change in the utility function. Thus, both $f(y^i)$ and $g(y^i)$ are information invariant.

Let us consider three classes of criteria for a player to change their tentative decision:

[4] If correlated strategies are taken into consideration, SIE of never nest response is equivalent to SIE of strictly dominated strategies

$\mathcal{C}_1 = \{K_i f(y^i) : i \in N, y^i \in X^i\}$.
$\mathcal{C}_2 = \{K_i(\vee_{y^i} f(y^i)) : i \in N, y^i \in X^i\}$.
$\mathcal{C}_3 = \{K_i(f(y^i) \vee g(y^i)) \wedge \neg K_i \neg f(y^i)) : i \in N, y^i \in X^i\}$.

Each of them represents condition which a player adopts as a criterion for him to change his tentative decision. A typical sentence, $K_i f(y^i)$, in $\mathcal{C}_1$ implies that i knows a specific alternative plan to increase his payoff, while a typical sentence, $K_i(\vee_{y^i} f(y^i))$, in $\mathcal{C}_2$ says that he merely knows of the existence of such a plan. Obviously, the conditions in $\mathcal{C}_1$ are stronger than those in $\mathcal{C}_2$. A formula in $\mathcal{C}_3$ means that i knows a specific alternative plan that does not decrease his payoff (i.e., $K_i(f(y^i) \vee g(y^i))$), and that it is possible that it will increase his payoff (i.e., $\neg K_i \neg f(y^i)$).

Benthem [2] considered iterative updatability of the conjunction of $\mathcal{C}_i^-$ ($i = 1, 2$), $\wedge_{\varphi \in \mathcal{C}_i}(\neg \varphi)$. He focused on a Kripke-model, in which any state is identified by a strategy profile chosen by the players, and every player knows only his own strategy. Then, he demonstrated the equivalence between the eliminated strategies and the states of the Kripke-model eliminated by information updates. According to his observation, the elimination of states involved by the announcement of a sentence in $\mathcal{C}_1$ (resp. $\mathcal{C}_2$) is equivalent to elimination of strictly (resp. weakly) dominated strategies. Similarly, we can easily show that the counterpart of $\mathcal{C}_3$ is weak domination.

Note that the dynamic information update process is applicable to any given epistemic state. Thus, the dynamic reformulation provides a solution concept for epistemically diverse situations, rather than an artificial situation to characterize existent solution concepts that Benthem [2] discussed.

4.3 Properties

Our first contribution here is to clarify the order independence of the elimination process from a dynamic viewpoint, which Benthem [2] did not address. It is well known that the set of SIE of strictly dominated strategies as well as the never best response strategies is uniquely determined while the algorithm above is non-deterministic. On the other hand, SIE of weakly dominated strategies is not. Gilboa, et al.[9] stated sufficient conditions for the order independence in non-epistemic terms.

According to Theorem 2, the order of information update does not matter at all, since all the sentences in $\mathcal{C}_1$ or $\mathcal{C}_2$ are information-monotonic. On the other hand, any sentence in $\mathcal{C}_3$ is information-monotonic, and thus, the set of SIE of weakly dominated strategies is order dependent. In summary, the well-known properties of order independence and dependence can be ascribed to information-monotonicity.

Second contribution is to clarify the role of information-monotonicity in comparison of the conditions generated by $\mathcal{C}_1$, $\mathcal{C}_2$, and $\mathcal{C}_3$. Obviously, $\mathcal{C}_1$ is stronger than $\mathcal{C}_2$ in the sense that $K_i f(y^i) \Rightarrow K_i(\vee_{y^i} f(y^i))$. It is, however, not trivial that the iterative updatability of $\mathcal{C}_1^-$ implies that of $\mathcal{C}_2^-$. Benthem [2] demonstrated this relation by a fixed-point method on Kripke-models. On the other hand, it is also a direct consequence of Comparison Theorem in the previous section since all sentence in $\mathcal{C}_1^-$ are monotonic. Further from Comparison theorem we obtain the same relation between $\mathcal{C}_1^-$ implies that of $\mathcal{C}_3^-$.

Third contribution is related to Nash equilibrium and the iterative updatability of $\mathcal{C}_i^-$. Consider the formula $q^* = \vee_{i \in N}(\vee_{y^i} f^i(y^i))$, which translates to 'the tentative decision is not a Nash equilibrium'. We focus on a sentence, $E = \neg(K_1 q^* \vee K_2 q^* \vee \cdots \vee K_n q^*)$, which can be translated into the following:

'No one knows that the tentative decision is not a Nash equilibrium'.

First, if E is true then it is iteratively updatable. Formally, we obtain the following lemma.

Lemma 1 (Idenpotency Lemma). *Assuming that* $q^* \in Q^*$*, if* $E = \neg(K_1 q^* \vee K_2 q^* \vee \cdots \vee K_n q^*)$*, then* $\vdash E \leftrightarrow \langle E \rangle E$.

Proof. $\vdash \langle E \rangle E \Rightarrow E$ is trivial. We show that $\vdash E \Rightarrow \langle E \rangle E$. It suffices to show that $\vdash E \Rightarrow \langle E \rangle \neg K_i q^*$ for all k.

From invariance of q^*, $\vdash \langle E \rangle q^* \leftrightarrow E \wedge q^*$. Further, $\vdash E \Rightarrow q^*$ and, thus, $\vdash (\neg E \Rightarrow \langle E \rangle q^*) \Rightarrow q^*$. By A2 and R2 we have that $\vdash \neg K_i q^* \Rightarrow \neg K_i(\neg E \Rightarrow \langle E \rangle q^*)$. Then $\vdash E \Rightarrow \neg K_i(\neg E \Rightarrow \langle E \rangle q^*)$ since $\vdash E \Rightarrow \neg K_i q^*$. Further, $\vdash \langle E \rangle \neg K_i q^* \leftrightarrow E \wedge \neg K_i(\neg E \Rightarrow \langle E \rangle q^*)$ by P2. It follows that $\vdash E \Rightarrow \langle E \rangle \neg K_i q^*$. □

Further, E means the iterative updatability of $\mathcal{C}_k^-$ $(k = 1, 2)$. That is, if none knows that the tentative decision is not a Nash equilibrium, then every player who changes the tentative decision only if he knows that the deviation to another strategy increases his payoff does not do so regardless of iterative information updates.

Theorem 4. *Let* $q \in Q^*$ *and* $E := \neg(K_1 q^* \vee K_2 q^* \vee \cdots \vee K_n q^*)$*. Assume that for all* $k = 1, 2, \ldots, n$*, there exists some* $i \in N$ *such that* $\vdash \neg\varphi_k \Rightarrow K_i q$*. Then,* $\vdash E \Rightarrow \langle \varphi_1 \rangle \langle \varphi_2 \rangle \ldots \langle \varphi_{m-1} \rangle \varphi_m$.

Proof. $\neg E$ is in $\mathcal{M}$, and $\vdash E \Rightarrow \varphi_l$ for all l. Define F_k and G_k by $F_1 = E$; $F_k = \langle F_{k-1} \rangle E$; $G_1 = \varphi_1$; $G_k = \langle G_{k-1} \rangle \varphi_k$. Then by Theorem 3, $\vdash F_k \Rightarrow G_k$. Further, by the previous lemma, we have $\vdash E \leftrightarrow \langle E \rangle E$, and thus, $\vdash E \Rightarrow F_k$. Therefore, $\vdash E \Rightarrow G_k$. □

In sum, the notion, 'none knows that the tentative decision is not Nash equilibrium', is, in our view, a noteworthy concept for analyzing epistemically diverse situations.

4.4 Information-variant utilities

Our results in the previous subsection depend on the assumption that the players' utility functions are invariant. To see this, consider a two-person game model with two states, a and b, and a variant utility function. An accessible relation R_1 is defined by xR_1y for all $x, y \in \{a, b\}$, while R_2 is defined by xR_2x for all $x \in \{a, b\}$. Player 1 has two possible strategies, T and B, while player 2 has L and R.

The utilities are represented in four tables below. They depend on the pair of strategies, the state of Kripke-model, and the set of remained states. The upper two tables represent the utilities when there remains only one state while the lower two represent those when the all states remain. In each square, the lower left value represents the utility of player 1, and the upper right value represents that of player 2.

1 \ 2	L	R
T	2, 1	0, 0
B	0, 0	2, 1

state a in $\{a\}$

1 \ 2	L	R
T	2, 1	0, 2
B	10, 0	2, 1

state b in $\{b\}$

1	L	R
T	2, 1	0, 0
B	5, 0	2, 1

state a in $\{a, b\}$

1	L	R
T	2, 1	0, 2
B	5, 0	2, 1

state b in $\{a, b\}$

All tables differ from each other in the utility of player 1 when (B, L) is chosen. Then, player 1 enjoys 0 if the state is a and he knows it, while he enjoys 10 if the state is b and he know it. When he does not know which of a and b is the true state he evaluates his utility as 5, which is a result from the expectation with probability 1/2 for each state.

Suppose that $x = (T, L)$ is the tentative decision in any state. Let φ denote the statement '$u_2(T, R) > u_2(x)$', and let ψ denote '$u_1(B, L) > u_1(x)$'. Then, $K_2\varphi, K_1\psi \in \mathcal{C}$. Note that in this case $K_1\psi$ is not information-monotonic.

(L, T) is not maintained because in any state on $\{a, b\}$, $K_1\psi$ is true. In state a, however, both $\neg K_1\psi$ and $\neg K_2\varphi$ become true after $\neg K_2\varphi$ is publicly announced. That is, $\langle \neg K_2\varphi \rangle \langle \neg K_1\psi \rangle \ldots$ is true. According to the theorems in the previous section, this phenomenon would not be observable if ψ and φ were information-monotonic.

5 Conclusion

We have discussed the epistemic conditions for the stability of strategies in a situation in which each player chooses a tentative decision under iterative public announcements about the other players' choices. To analyze these conditions, we extended Plaza's public announcement logic by adding the notions of information-invariance and information-monotonicity. By means of these notions, we clarified the conditions for robustness with respect to the order of information update, that was not investigated in [2].

Our analysis has room for improvement. The applications presented in this paper are still simple. Moreover, we focused on the case in which any information update is done through public announcement. That is, we did not consider various types of update process as is discussed in [4]. Our simple logic, however, might open up new approaches to related research issues such as syntactic analysis.

Acknowledgement

We thank Thomas Ågotnes, Nobu-Yuki Suzuki, and Mamoru Kaneko for their useful comments on this work.

References

1. J. van Behthem, J. van Eijck, and B. Kooi. Logics of Communication and Change. to appear in *Information and Computation.*
2. J. van Benthem. Rational Dynamics and Epistemic Logic in Games. *International Game Theory Review*, vol.9, pp.13-45, 2007.
3. D. Bernheim. Rationalizable Strategic Behavior. *Econometrica*, vol.52, pp.1007-1028, 1984.
4. H.P. van der Ditmarsch, W. van der Hoek, and W. B. Kooi. *Dynamic Epistemic Logic*, Springer, 2007.
5. R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning About Knowledge.* MIT press, 1995.
6. D. Gale, H.W. Kuhn, and A.W. Tucker. Reduction of game matrices, *Lectures on the Theory of Games, Annals of Mathematics Studies*, No. 37, edited by H.W. Kuhn and A.W. Tucker, Princeton university press, Princeton, 89–96, 1950.
7. D. Gale. A theory of n-person game with perfect information. *Proceedings of the National Academy of Sciences of the United States of America*, vol 39, Issue 6, pp. 496-501, 1953.
8. J. Gerbrandy. Dynamic epistemic logic. *Logic, Language and Computation*, vol.2, pages 67-84, 1999.
9. I. Gilboa, E.Kalai, and E. Zemel. On the order of eliminating dominated strategies *Operations Research Letters 9*, vol 9, pp.85-89, 1990.
10. J.C. Harsanyi. Games with Incomplete Information Played by Bayesian Players: Part II Bayesian Equilibrium Points, *Management Science*, vol.14, pp. 320-334, 1968.
11. J. Hintikka. *Knowledge and Belief*, Cornell University Press, 1962.

12. G. E. Hughes and M. J. Creswell. *A New Introduction to Modal Logic*, Routledge, 1996.
13. R.D. Luce and H. Raiffa. *Games and Decisions: Introduction and Critical Survey* Wiley: New York, 1957.
14. D. Pearce. Rationalizable Strategic Behavior and the Problem of Perfection. *Econometrica*, vol.52, pp.1029-1050, 1984.
15. J. Plaza. Logic of Public Communications. *4th Int'l Symposium on Methodologies for Intelligent Systems*, pp.201-216, 1989. (reprinted version is in *Synthese*, vol.158, pp.165-179, 2007.)
16. L. Samuelson. Dominated strategies and common knowledge, *Games and Economic Behavior*, vol. 4(2), pp.284-313,1992.
17. T.C. Tan, and S.R. Werlang. The Bayesian foundations of solution concepts of games, *Journal of Economic Theory*, vol.45(2), pp. 370-391, 1988.