Title	Analyzing Reliability Change in Legal Case
Author(s)	Jirakunkanok, Pimolluck; Sano, Katsuhiko; Tojo, Satoshi
Citation	Lecture Notes in Computer Science, 9067: 274-290
Issue Date	2015-08-25
Type	Journal Article
Text version	author
URL	http://hdl.handle.net/10119/14212
Rights	This is the author-created version of Springer, Pimolluck Jirakunkanok, Katsuhiko Sano, Satoshi Tojo, Lecture Notes in Computer Science, 9067, 2015, 274-290. The original publication is available at www.springerlink.com, http://dx.doi.org/10.1007/978-3-662-48119-6_20
Description	JSAI-isAI 2014 Workshops, LENLS, JURISIN, and GABA, Kanagawa, Japan, October 27-28, 2014, Revised Selected Papers

Analyzing Reliability Change in Legal Case

Pimolluck Jirakunkanok, Katsuhiko Sano, and Satoshi Tojo

School of Information Science,
Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
{pimolluck.jira,v-sano,tojo}@jaist.ac.jp

Abstract. A consideration of the reliability plays a significant role in agent communication. An agent can change her belief about the reliability ordering between the other agents with respect to new incoming information. In order to analyze reliability change of an agent, this paper proposes a logical formalization with two dynamic operators, i.e., downgrade and upgrade operators. The downgrade operator allows an agent to downgrade some specified agents to be less reliable corresponding to the degree of reliability, while the upgrade operator allows an agent to upgrade them to be more reliable. Furthermore, we demonstrate our formalization by a legal case from Thailand.

Keywords: reliability change, belief, legal case, modal logic, signed information

1 Introduction

In agent communication, an agent needs some criteria to decide which information she should believe. A common criterion is to consider the reliability of an information source. If the agent considers that a source of received information is reliable, she would accept and might believe the received information. On the other hand, the agent may reject the received information if she considers that the source is not reliable. Legal proceedings are a typical example of agent communication that the reliability has a strong influence on a judge's decision in a court. Thus, the judge needs the reliability, i.e., when a judge receives a piece of information from a witness, the judge should consider if the witness is reliable or not. In addition, when the judge receives new information, she might change her belief about the reliability of the witness. This paper aims to investigate an effect of reliability change of the judge in legal judgment.

Recently, many studies [1–3] presented the use of logic-based approaches in the legal systems. Dynamic epistemic logic (DEL) [4, 5] is a logical tool to study reasoning about information change due to communication between agents. With these principles, several works [6–8] proposed a logical framework for formalizing the reliability. Among of them, Lorini et al. [8] introduced a modal framework for reasoning about signed information. In their framework, the agents can keep track of the information source by using the notion of signed statement. They also considered the notion of reliability over the information sources. However, they did not deal with the dynamics of the reliability relation of agents.

For this reason, we propose to formalize reliability change of an agent. First, we apply a concept of signed statement based on [8] to formalize the source of information. Then, we introduce two dynamic operators, i.e. downgrade and upgrade operators, in order to capture the change of reliability ordering between agents. The downgrade operator is used for downgrading the reliability of agents, while the upgrade operator is used for upgrading. Finally, we reformulate a careful policy [8] in terms of DEL and employ it to consider which pieces of received signed information an agent should believe. Moreover, we demonstrate our formalization in an example of a legal case from Thailand.

The remainder of this paper is organized as follows. Section 2 describes the target legal case. Then, a formal tool for analyzing the legal case is presented in Section 3. In Section 4, we propose a dynamic logical analysis of the target legal case. Finally, our conclusion and future works are stated in Section 5.

2 Target Legal Case

Firstly, we summarize a story of our target legal case that occurred on 26th January 2003 in Trang province, Thailand [1] as follows:

> One day, a victim v had a drink with his friends f_1, f_2 and d at f_2's house. After that, v was punched and stabbed with a hand scraper in the back by an offender, and as a result, v has bleeding in the lung. However, v was still alive.

In the inquiry stage, a police po, who is an inquiry official, interviewed four witnesses v, f_1, f_2, mo that gave the following statements.

(I_1) v tells that d is the offender who punched and stabbed v.
(I_2) f_1 also tells that d is the offender who punched and stabbed v.
(I_3) f_2 states that v and d had a dispute, but did not have any fighting.
(I_4) mo, who is v's mother, tells that d is the offender according to v's saying. More details can be shown as follows:

> At night of the accident, mo visited v in the hospital. Then, v told her that v went to have a drink with d, f_1 and f_2 at f_2's house. During drinking, v and d had a dispute, then d punched v and stabbed with a hand scraper in the back of v.

From the interview, po accused d of attempting to kill v.

In the Civil Court, v and f_1 changed their statements as follows:

v tells that one of a group of unknown teenagers is the offender who punched and stabbed v by a knife. More details can be shown as follows:

[1] This legal case can be referred from `http://deka2007.supremecourt.or.th/deka/web/search.jsp` (in Thai).

> At 19 o'clock, v and f_1 were invited to drink by x who was their neighbor. After drinking, v and f_1 went to a market. While f_1 was riding a motorcycle from x's house, a group of unknown teenagers came to punch v. Then, one of them armed with a knife stabbed in the back of v.

f_1 can only state that v was punched by d, but cannot state that v was stabbed by d or not. More details can be shown as follows:

> At 18 o'clock, v and f_1 were invited to drink by d. Then, v and f_1 went to f_2's house by a motorcycle (v was a rider), and d also followed them. Next, v had a drink with f_1, f_2, d and two other friends at f_2's house. Around 21 o'clock, v and d had a dispute and then d punched v. f_2 came to forbid them from fighting, while f_1 went to bring the motorcycle. After that, v came to sit behind f_1's motorcycle and said that he was stabbed.

Moreover, *po* was called to be a witness for testifying all statements in the inquiry stage. Thus, there are six testimonies in the Civil Court as follows:

(T_1) v tells that one of a group of unknown teenagers is the offender who punched and stabbed v by a knife.
(T_2) f_1 can only state that v was punched by d, but cannot state that v was stabbed by d or not.
(T_3) *po* states that v tells that d is the offender who punched and stabbed v.
(T_4) *po* states that f_1 tells that d is the offender who punched and stabbed v.
(T_5) *po* states that f_2 states that v and d had a dispute, but did not have any fighting.
(T_6) *po* states that *mo* tells that d is the offender according to v's saying.

From the above testimonies, testimonies of v and f_1 in the inquiry stage (T_3 and T_4) are more reliable than that in the Civil Court (T_1 and T_2) because of the following reasons. First, the judge believed that *po* and f_2 had never had any arguments against d. So, there is no reason that they will allege or testify against d to be punished. Second, according to T_1 and T_2, the judge believed that v and f_1 tried to distort the facts in order to prevent d who is their friend from the punishment. Therefore, the judge decided that d was the offender and intended to kill v by the following reasons.

- Since the hand scrapper is a dangerous weapon, d uses it in a possibly lethal attack. This shows that d intends to kill v.
- d stabs v while v is turning back. At that time, d can choose other alternative positions for attacking. Nevertheless, d strongly stabs v in the lung that is a vital organ. It is obvious that d intends to kill v.
- From the statement of the doctor, v is seriously injured, i.e., there is air leaking and bleeding in the chest cavity and the lung, and would be dead unless v gets the treatment in time. This shows that the attack of d is possibly lethal.

For this reason, the Civil Court judged d to be sentenced to ten years' imprisonment by Article 288 and Article 80 of Penal Code: [2]

[2] An English translation of articles can be referred from `http://www.thailaws.com/`.

> *Article 288* (offence causing death): Whoever, murdering the other person, shall be punished by death or imprisoned as from fifteen years to twenty years.

> *Article 80* (commitment): Whoever commences to commit an offence, but does not carry it through, or carries it through, but does not achieve its end, is said to attempt to commit an offence. Whoever attempts to commit an offence shall be liable to two-thirds of the punishment as provided by the law for such offence.

In the Appeal Court and the Supreme Court, d appealed that he did not intend to kill v; in fact, he only intended to attack v. However, the judge agreed with the decision of the Civil Court and adopted the result, i.e., d is imprisoned for ten years by Articles 288 and Article 80 of Penal Code.

3 Formal Tool for Analyzing Target Legal Case

3.1 Static Logic of Agents' Beliefs for Signed Information

To analyze the previous legal case from a logical point of view, we introduce a modal language, based on previous work [8], which enables us to formalize the agent's belief, the reliability of information sources, and signed information.

Let G be a fixed *finite* set of agents. Our syntax $\mathcal{L}$ consists of the following vocabulary: (i) a countably infinite set $\mathsf{Prop} = \{ p, q, r, ... \}$ of propositional letters, (ii) Boolean connectives: $\neg$, $\wedge$, (iii) the belief operators $\mathsf{Bel}(a, \cdot)$ $(a \in G)$, (iv) the signature operators $\mathsf{Sign}(a, \cdot)$ $(a \in G)$, and (v) the constants for reliability ordering $b \leqslant_a c$ $(a, b, c \in G)$. A set of formulas of $\mathcal{L}$ is inductively defined as follows:

$$\varphi ::= p \,|\, \neg\varphi \,|\, \varphi \wedge \varphi \,|\, \mathsf{Bel}(a, \varphi) \,|\, \mathsf{Sign}(a, \varphi) \,|\, b \leqslant_a c,$$

where $p \in \mathsf{Prop}$ and $a, b, c \in G$. For intuitive readings of formulas, the reader can refer to Table 1. Note that $b <_a c$ stands for b is strictly more reliable than c, i.e., $(b \leqslant_a c) \wedge \neg(c \leqslant_a b)$, and $b \approx_a c$ which stands for b and c are equally reliable can be defined as $(b \leqslant_a c) \wedge (c \leqslant_a b)$. We define $\vee$, $\rightarrow$, $\leftrightarrow$ as ordinary abbreviations. Our syntax is different from [8] in at least two respects. First, we do not introduce the universal quantifier for agents. This is because we realized that most of the ideas in [8] are done without quantifiers for agents when the set of agents is finite, i.e., the universal quantifier for a finite domain is just reduced to the conjunction of finite conjuncts. Second, we relativize the notion of reliability ordering $\leqslant$ to each agent. In order to analyze our example from a logical perspective, we need to formalize belief change of a judge of the Civil Court and we regard that belief change is induced by reliability change. However, there is no need for us to change the reliability ordering of the other agents other than the judge of the Civil Court. This is why we propose the notion of reliability ordering between agents depending on a particular agent's perspective.

$\mathsf{Bel}(a, \varphi)$	: agent a believes that φ.
$\mathsf{Sign}(a, \varphi)$	: agent a signs statement φ.
$b \leqslant_a c$	: from agent a's perspective, agent b is at least as reliable as agent c.
$\mathsf{Sign}\big(a, \mathsf{Sign}(b, \varphi)\big)$	: agent a signs statement that agent b signs statement φ.
$\mathsf{Bel}\big(a, \mathsf{Sign}(b, \varphi)\big)$	: agent a believes that agent b signs statement φ.
$\mathsf{Bel}\big(a, b \leqslant_a c\big)$	: agent a believes that from agent a's perspective, agent b is at least as reliable as agent c.

Table 1. Examples of Static Logical Formalization

Let us provide Kripke semantics for our syntax. A *model* $\mathfrak{M}$ is a tuple

$$\mathfrak{M} = (W, (R_a)_{a \in G}, (S_a)_{a \in G}, (\preccurlyeq_a)_{a \in G}, V),$$

where W is a non-empty set of states, called *domain*, $R_a \subseteq W \times W$ is an accessibility relation representing beliefs, $S_a \subseteq W \times W$ is an accessibility relation representing signatures, $\preccurlyeq_a$ is a function which maps from W to $\mathcal{P}(G \times G)$ representing the reliability orderings between agents corresponding to each agent, and $V : \mathsf{Prop} \to \mathcal{P}(W)$ is a valuation. In what follows, we simply write $b \preccurlyeq_a^w c$ for $(b, c) \in \preccurlyeq_a (w)$. For any binary relation X on W and any state $w \in W$, we write $X(w)$ to mean $\{ v \in W \mid (w, v) \in X \}$.

Given any model $\mathfrak{M}$, any state $w \in W$, and any formula φ, we define the *satisfaction relation* $\mathfrak{M}, w \models \varphi$ inductively as follows:

$$\begin{array}{lll} \mathfrak{M}, w \models p & \text{iff} & w \in V(p) \\ \mathfrak{M}, w \models \neg\varphi & \text{iff} & \mathfrak{M}, w \not\models \varphi \\ \mathfrak{M}, w \models \varphi \wedge \psi & \text{iff} & \mathfrak{M}, w \models \varphi \text{ and } \mathfrak{M}, w \models \psi \\ \mathfrak{M}, w \models b \leqslant_a c & \text{iff} & b \preccurlyeq_a^w c \\ \mathfrak{M}, w \models \mathsf{Sign}(a, \varphi) & \text{iff} & \mathfrak{M}, v \models \varphi \text{ for all states } v \text{ such that } wS_a v \\ \mathfrak{M}, w \models \mathsf{Bel}(a, \varphi) & \text{iff} & \mathfrak{M}, v \models \varphi \text{ for all states } v \text{ such that } wR_a v \end{array}$$

A formula φ is *valid* in a model $\mathfrak{M}$ if $\mathfrak{M}, w \models \varphi$ for all states w of $\mathfrak{M}$.

Definition 1. *A model* $\mathfrak{M} = (W, (R_a)_{a \in G}, (S_a)_{a \in G}, (\preccurlyeq_a)_{a \in G}, V)$ *is a si-model* (*a model for signed information*) *if the following conditions are satisfied:*

(i) R_a *is transitive* (wRv *and* vRu *jointly imply* wRu *for all states* w, v, u) *and Euclidean* (wRv *and* wRu *jointly imply* vRu *for all states* w, v, u).
(ii) S_a *is serial* (*for any state* w, *there is some state* v *such that* $wS_a v$), *transitive and Euclidean.*
(iii) $\preccurlyeq_a^w \subseteq G \times G$ *is a total pre-ordering between agents, i.e.,* $\preccurlyeq_a^w$ *is reflexive* ($b \preccurlyeq_a^w b$ *for all agents* b), *transitive, and comparable* (*for any agents* b *and* c, $b \preccurlyeq_a^w c$ *or* $c \preccurlyeq_a^w b$).

The first and second items of this definition ensure us that we never sign a contradiction (due to seriality of S_a), and $\mathsf{Bel}(a, \cdot)$ and $\mathsf{Sign}(a, \cdot)$ are both positively and negatively introspective. Corresponding to these constraints, we easily obtain the following validities.

Proposition 1. *The following are valid in all si-models: for all a, b, $c \in G$,*

(i) $\mathsf{Bel}(a, p) \rightarrow \mathsf{Bel}(a, \mathsf{Bel}(a, p))$ *and* $\neg\mathsf{Bel}(a, p) \rightarrow \mathsf{Bel}(a, \neg\mathsf{Bel}(a, p))$.
(ii) $\neg\mathsf{Sign}(a, \bot)$, $\mathsf{Sign}(a, p) \rightarrow \mathsf{Sign}(a, \mathsf{Sign}(a, p))$, *and* $\neg\mathsf{Sign}(a, p) \rightarrow \mathsf{Sign}(a, \neg\mathsf{Sign}(a, p))$.
(iii) $b \leqslant_a b$, $(b \leqslant_a c \wedge c \leqslant_a d) \rightarrow b \leqslant_a d$, *and* $b \leqslant_a c \vee c \leqslant_a b$.

Based on Definition 1 and the idea of [8], we can rank agents by giving a partition $(\mathsf{C}^a_i)_{i \leq M}$ to G, where M is a natural number representing the maximum rank (such M always exists because G is finite) and we read $c \in \mathsf{C}^a_i$ as 'from agent a's viewpoint, the rank of agent c is i'. As a result, the agents who are equally reliable are categorized in the same group. C^a_1 which stands for 'a group of agents which is the most reliable from a's perspective' can be defined by the following formula:

$$c \in \mathsf{C}^a_1 =_{def} \bigwedge\nolimits_{b \in G} (c \leqslant_a b),$$

where we recall that G is a finite set of agents and $a, b, c \in G$. Note that we relativize the notion C^a_i to a specified agent a because the notion of reliability ordering $\leqslant_a$ depends on a specified agent a. This is a difference from [8] because [8] did not consider C_i depending on a specified agent. Then, we can rank the group of agents C^a_i such that $i > 1$ as follows:

$$c \in \mathsf{C}^a_i =_{def} \left(\left(\bigwedge_{1 \leq j \leq i-1} \neg(c \in \mathsf{C}^a_j) \right) \wedge \left(\bigwedge_{b \in G} \left(\left(\bigwedge_{1 \leq j \leq i-1} \neg(b \in \mathsf{C}^a_j) \right) \rightarrow (c \leqslant_a b) \right) \right) \right).$$

This implies that all agents in C^a_i are equally reliable, and if $i <_{\mathbb{N}} j$ then $c <_a b$ for all agents $c \in \mathsf{C}^a_i$ and agent $b \in \mathsf{C}^a_j$.

Theorem 1. *The set of all valid formulas on all si-models is axiomatized by:*

- *all propositional tautologies*
- $\mathsf{Bel}(a, p \rightarrow q) \rightarrow (\mathsf{Bel}(a, p) \rightarrow \mathsf{Bel}(a, q))$ $(a \in G)$
- $\mathsf{Sign}(a, p \rightarrow q) \rightarrow (\mathsf{Sign}(a, p) \rightarrow \mathsf{Sign}(a, q))$ $(a \in G)$
- *From* φ *we may infer* $\mathsf{Bel}(a, \varphi)$ $(a \in G)$
- *From* φ *we may infer* $\mathsf{Sign}(a, \varphi)$ $(a \in G)$
- *uniform substitution and modus ponens,*

as well as all listed formulas of Proposition 1.

3.2 Downgrade and Upgrade Operations for Agents

In order to change a reliability ordering between agents from a particular agent's perspective, we introduce two dynamic operators, i.e., the downgrade operator $[H \Downarrow^a_\varphi]$ and the upgrade operator $[H \Uparrow^a_\varphi]$, where $H \subseteq G$ is a set of agents. Our intended reading of $[H \Downarrow^a_\varphi]\psi$ is 'after the agent a downgraded the agents in H who sign the statement φ, ψ holds', and we can read $[H \Uparrow^a_\varphi]\psi$ as 'after the agent a upgraded the agents in H who sign the statement φ, ψ holds'. Semantically speaking, $[H \Downarrow^a_\varphi]$ makes all agents in H who sign φ less reliable than all the

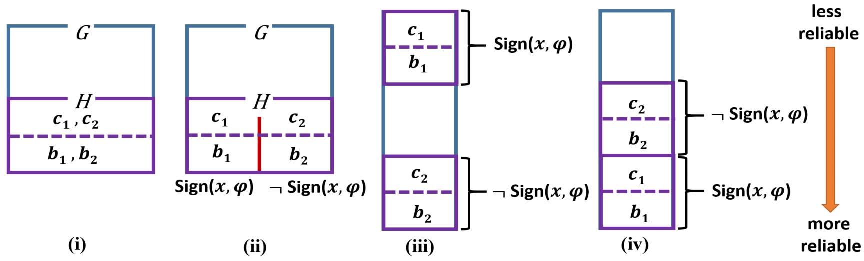

Fig. 1. Downgrading and Upgrading. (iii) is an effect of downgrading $[H \Downarrow^a_\varphi]$ to (ii), and (iv) is an effect of upgrading $[H \Uparrow^a_\varphi]$ to (ii).

other agents, and $[H \Uparrow^a_\varphi]$ makes all agents in H who sign φ more reliable than all the other agents.

Before giving a detailed semantics, let us demonstrate the effect of $[H \Downarrow^a_\varphi]$ and $[H \Uparrow^a_\varphi]$ by figures. Firstly, we assume that a rectangle G of Fig. 1(i) represents a fixed finite set of agents. Secondly, we will select a specified set of agents in order to change their reliability ordering that can be represented by a rectangle H, and we assume that $b_1 \approx_a b_2 <_a c_1 \approx_a c_2$ holds, i.e., agents b_1 and b_2 which are equally reliable are more reliable than agents c_1 and c_2 which are equally reliable from agent a's perspective. In this sense, b_1, b_2, c_1 and c_2 are situated as in Fig. 1(i). Then, if we focus on the agents who sign the statement φ, H is divided into two equal vertical parts, i.e., $\mathsf{Sign}(x, \varphi)$ and $\neg\mathsf{Sign}(x, \varphi)$ as in Fig. 1(ii), namely by the set $\{x \in H \mid \mathfrak{M}, w \models \mathsf{Sign}(x, \varphi)\}$ and the set $\{x \in H \mid \mathfrak{M}, w \models \neg\mathsf{Sign}(x, \varphi)\}$. Next, if the agent a downgrades all the agents signing the statement φ, we downgrade all of them less reliable than the other agents as in Fig.1(iii). On the other hand, if the agent a upgrades all the agents signing the statement φ, we upgrade all of them more reliable than the other agents as in Fig.1(iv).

Definition 2. *Given a Kripke model* $\mathfrak{M} = (W, (R_a)_{a\in G}, (S_a)_{a\in G}, (\preccurlyeq_d)_{d\in G}, V)$, *a semantic clause for* $[H \Downarrow^a_\varphi]$ *on* $\mathfrak{M}$ *and* $w \in W$ *is defined by:*

$$\mathfrak{M}, w \models [H \Downarrow^a_\varphi]\psi \ \text{ iff } \ \mathfrak{M}^{H\Downarrow^a_\varphi}, w \models \psi,$$

where $\mathfrak{M}^{H\Downarrow^a_\varphi} = (W, (R_a)_{a\in G}, (S_a)_{a\in G}, (\preccurlyeq'_d)_{d\in G}, V)$ *and* $\preccurlyeq'_d$ *is defined as: for all* $u \in W$*:*

- *if* $d \neq a$, *we put* $\preccurlyeq'^u_d = \preccurlyeq^u_d$.
- *otherwise (if* $d = a$*), we define* $b \preccurlyeq'^u_a c$ *iff*

$$\big(b, c \in H \text{ and } \mathfrak{M}, u \models \mathsf{Sign}(b, \varphi) \wedge \mathsf{Sign}(c, \varphi) \text{ and } b \preccurlyeq^u_a c\big) \text{ or}$$
$$\big(b, c \in (G \setminus H) \cup \{x \in H \mid \mathfrak{M}, u \models \neg\mathsf{Sign}(x, \varphi)\} \text{ and } b \preccurlyeq^u_a c\big) \text{ or}$$
$$\big(b \in (G \setminus H) \cup \{x \in H \mid \mathfrak{M}, u \models \neg\mathsf{Sign}(x, \varphi)\} \text{ and } c \in H \text{ and } \mathfrak{M}, u \models \mathsf{Sign}(c, \varphi)\big)^3$$

[3] In this case, since there is no relation between agents b and c, $b \preccurlyeq^u_a c$ is omitted.

Definition 3. *Given a Kripke model* $\mathfrak{M} = (W, (R_a)_{a\in G}, (S_a)_{a\in G}, (\preccurlyeq_d)_{d\in G}, V)$, *a semantic clause for* $[H \Uparrow^a_\varphi]$ *on* $\mathfrak{M}$ *and* $w \in W$ *is defined by:*

$$\mathfrak{M}, w \models [H \Uparrow^a_\varphi]\psi \text{ iff } \mathfrak{M}^{H\Uparrow^a_\varphi}, w \models \psi,$$

where $\mathfrak{M}^{H\Uparrow^a_\varphi} = (W, (R_a)_{a\in G}, (S_a)_{a\in G}, (\preccurlyeq'_d)_{d\in G}, V)$ *and* $\preccurlyeq'_d$ *is defined as: for all* $u \in W$*:*

- *if* $d \neq a$*, we put* $\preccurlyeq'^u_d = \preccurlyeq^u_d$.
- *otherwise (if* $d = a$*), we define* $b \preccurlyeq'^u_a c$ *iff*

$$\big(b, c \in H \text{ and } \mathfrak{M}, u \models \mathsf{Sign}(b,\varphi) \wedge \mathsf{Sign}(c,\varphi) \text{ and } b \preccurlyeq^u_a c\big) \text{ or}$$
$$\big(b, c \in (G \setminus H) \cup \{x \in H \mid \mathfrak{M}, u \models \neg\mathsf{Sign}(x,\varphi)\} \text{ and } b \preccurlyeq^u_a c\big) \text{ or}$$
$$\big(c \in (G \setminus H) \cup \{x \in H \mid \mathfrak{M}, u \models \neg\mathsf{Sign}(x,\varphi)\} \text{ and } b \in H \text{ and } \mathfrak{M}, u \models \mathsf{Sign}(b,\varphi)\big)^3$$

Proposition 2. *If* $\mathfrak{M}$ *is a si-model, then both* $\mathfrak{M}^{H\Uparrow^a_\varphi}$ *and* $\mathfrak{M}^{H\Downarrow^a_\varphi}$ *are si-models.*

Proposition 3 (Recursive Validities). *The following are valid on all models. Moreover, if* ψ *is valid on all models, then* $[H \Downarrow^a_\varphi]\psi$ *is also valid on all models.*

$$\begin{array}{lcll}
[H \Downarrow^a_\varphi]p & \leftrightarrow & p & \\
[H \Downarrow^a_\varphi](b \leqslant_d c) & \leftrightarrow & b \leqslant_d c & (d \neq a) \\
[H \Downarrow^a_\varphi](b \leqslant_a c) & \leftrightarrow & b \leqslant_a c & (b, c \in G \setminus H) \\
[H \Downarrow^a_\varphi](b \leqslant_a c) & \leftrightarrow & \big(\mathsf{Sign}(b,\varphi) \wedge \mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c)\big) \vee & \\
 & & \big(\neg\mathsf{Sign}(b,\varphi) \wedge \neg\mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c)\big) \vee & \\
 & & \big(\neg\mathsf{Sign}(b,\varphi) \wedge \mathsf{Sign}(c,\varphi)\big) & (b, c \in H) \\
[H \Downarrow^a_\varphi](b \leqslant_a c) & \leftrightarrow & \mathsf{Sign}(c,\varphi) \vee \big(\neg\mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c)\big) & (c \in H, b \in G \setminus H) \\
[H \Downarrow^a_\varphi](b \leqslant_a c) & \leftrightarrow & \neg\mathsf{Sign}(b,\varphi) \wedge (b \leqslant_a c) & (b \in H, c \in G \setminus H) \\
[H \Downarrow^a_\varphi]\neg\psi & \leftrightarrow & \neg[H \Downarrow^a_\varphi]\psi & \\
[H \Downarrow^a_\varphi](\psi_1 \wedge \psi_2) & \leftrightarrow & [H \Downarrow^a_\varphi]\psi_1 \wedge [H \Downarrow^a_\varphi]\psi_2 & \\
[H \Downarrow^a_\varphi]\mathsf{Sign}(b,\psi) & \leftrightarrow & \mathsf{Sign}(b, [H \Downarrow^a_\varphi]\psi) & \\
[H \Downarrow^a_\varphi]\mathsf{Bel}(b,\psi) & \leftrightarrow & \mathsf{Bel}(b, [H \Downarrow^a_\varphi]\psi) &
\end{array}$$

Proposition 4 (Recursive Validities). *The following are valid on all models. Moreover, if* ψ *is valid on all models, then* $[H \Uparrow^a_\varphi]\psi$ *is also valid on all models.*

$$\begin{array}{lcll}
[H \Uparrow^a_\varphi]p & \leftrightarrow & p & \\
[H \Uparrow^a_\varphi](b \leqslant_d c) & \leftrightarrow & b \leqslant_d c & (d \neq a) \\
[H \Uparrow^a_\varphi](b \leqslant_a c) & \leftrightarrow & b \leqslant_a c & (b, c \in G \setminus H) \\
[H \Uparrow^a_\varphi](b \leqslant_a c) & \leftrightarrow & \big(\mathsf{Sign}(b,\varphi) \wedge \mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c)\big) \vee & \\
 & & \big(\neg\mathsf{Sign}(b,\varphi) \wedge \neg\mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c)\big) \vee & \\
 & & \big(\mathsf{Sign}(b,\varphi) \wedge \neg\mathsf{Sign}(c,\varphi)\big) & (b, c \in H) \\
[H \Uparrow^a_\varphi](b \leqslant_a c) & \leftrightarrow & \neg\mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c) & (c \in H, b \in G \setminus H) \\
[H \Uparrow^a_\varphi](b \leqslant_a c) & \leftrightarrow & \mathsf{Sign}(b,\varphi) \vee \big(\neg\mathsf{Sign}(b,\varphi) \wedge (b \leqslant_a c)\big) & (b \in H, c \in G \setminus H) \\
[H \Uparrow^a_\varphi]\neg\psi & \leftrightarrow & \neg[H \Uparrow^a_\varphi]\psi & \\
[H \Uparrow^a_\varphi](\psi_1 \wedge \psi_2) & \leftrightarrow & [H \Uparrow^a_\varphi]\psi_1 \wedge [H \Uparrow^a_\varphi]\psi_2 & \\
[H \Uparrow^a_\varphi]\mathsf{Sign}(b,\psi) & \leftrightarrow & \mathsf{Sign}(b, [H \Uparrow^a_\varphi]\psi) & \\
[H \Uparrow^a_\varphi]\mathsf{Bel}(b,\psi) & \leftrightarrow & \mathsf{Bel}(b, [H \Uparrow^a_\varphi]\psi) &
\end{array}$$

3.3 Private Announcements

This section introduces a new dynamic operator for private announcement $[\varphi \rightsquigarrow a]$ (whose reading is "after a private announcement of φ to agent a"), where the idea of privateness is realized by the property that the other agent than a will not notice the recipient a's belief change. One of the merits of this operator is that we do not specify information of sender of message φ, while agent a is a recipient of the message. This means that we may use this operator also for self-decision of agent a, i.e., the sender and the recipient are the same. This section demonstrates that $[\varphi \rightsquigarrow a]$ can capture both (i) the *tell*-action $[\mathsf{Tell}(b, a, \varphi)]$ from [8]: 'agent b tells to agent a that a certain statement φ is true' and (ii) one of aggregation policies from [8] called the *careful policy*'. We note that Lorini et al. [8] did not propose a logical treatment from dynamic epistemic viewpoints for any of aggregation policies. Moreover, we note that the sender and the recipient are regarded as the same to capture the careful policy by our new operator $[\varphi \rightsquigarrow a]$.

Action Model for Private Announcements In order to capture this private action, we introduce the following special structure (called *action model* in dynamic epistemic logic, the reader may find a similar structure in [9, 4]).

Definition 4. *The action model for private announcements of φ to agent a is a tuple $(E, (D_c)_{c\in G}, (U_a)_{a\in G}, \mathrm{pre})$ such that E consists of two actions: φ-announcing action $!_\varphi$ to agent a and non-announcing action $\top$, and $D_a = \{(!_\varphi, !_\varphi), (\top, \top)\}$ and $D_c = \{(!_\varphi, \top), (\top, \top)\}$ if $c \neq a$, $U_c = \{(!_\varphi, \top), (\top, \top)\}$ for all $c \in G$, and pre assigns a precondition to each action, i.e., $\mathrm{pre}(!_\varphi) = \varphi$ and $\mathrm{pre}(\top) = \top$.*

Definition 5. *Given a Kripke model $\mathfrak{M} = (W, (R_c)_{c\in G}, (S_c)_{c\in G}, (\preccurlyeq_c)_{c\in G}, V)$, a semantic clause for $[\varphi \rightsquigarrow a]\psi$ on $\mathfrak{M}$ and $w \in W$ is defined as follows:*

$$\mathfrak{M}, w \models [\varphi \rightsquigarrow a]\psi \text{ iff } \mathfrak{M}^{\varphi\rightsquigarrow a}, (w, !_\varphi) \models \psi,$$

where $\mathfrak{M}^{\varphi\rightsquigarrow a} = (W', (R'_c)_{c\in G}, (S'_c)_{c\in G}, (\preccurlyeq'_c)_{c\in G}, V')$ is the updated model by the action model of Definition 4, i.e.,

- *$W' := W \times E = W \times \{!_\varphi, \top\}$.*
- *$(w, e)R'_c(v, f)$ iff wR_cv and $(e, f) \in D_c$ and $\mathfrak{M}, v \models \mathrm{pre}(f)$ (for all $c \in G$).*
- *$(w, e)S'_c(v, f)$ iff wS_cv and $(e, f) \in U_c$ (for all $c \in G$).*
- *$d \preccurlyeq_c^{\prime(w,e)} d'$ iff $d \preccurlyeq_c^w d'$.*
- *$(w, e) \in V'(p)$ iff $w \in V(p)$.*

Proposition 5. *If $\mathfrak{M}$ is a si-model, then $\mathfrak{M}^{\varphi\rightsquigarrow a}$ is also a si-model.*

Proposition 6 (Recursive Validities). *The following are valid on all models. Moreover, if ψ is valid on all models, then $[\varphi \rightsquigarrow a]\psi$ is also valid on all models.*

$$\begin{array}{ll}
[\varphi \rightsquigarrow a]p & \leftrightarrow p \\
[\varphi \rightsquigarrow a]d \leqslant_c d' & \leftrightarrow d \leqslant_c d' \\
[\varphi \rightsquigarrow a]\neg\varphi & \leftrightarrow \neg[\varphi \rightsquigarrow a]\varphi \\
[\varphi \rightsquigarrow a](\psi \wedge \theta) & \leftrightarrow [\varphi \rightsquigarrow a]\psi \wedge [\varphi \rightsquigarrow a]\theta \\
[\varphi \rightsquigarrow a]\mathsf{Bel}(a,\psi) & \leftrightarrow \mathsf{Bel}(a, \varphi \rightarrow [\varphi \rightsquigarrow a]\psi) \\
[\varphi \rightsquigarrow a]\mathsf{Bel}(c,\psi) & \leftrightarrow \mathsf{Bel}(c,\psi) \quad\quad (a \neq c) \\
[\varphi \rightsquigarrow a]\mathsf{Sign}(c,\psi) & \leftrightarrow \mathsf{Sign}(c,\psi)
\end{array}$$

Note that the axiom $[\varphi \rightsquigarrow a]\mathsf{Bel}(c,\psi) \leftrightarrow \mathsf{Bel}(c,\psi)$ captures that the action of a's privately receiving message φ will not affect of the other agents' beliefs than a.

Theorem 2. *The set of all valid formulas of the expanded syntax of $\mathcal{L}$ with $[H \Downarrow^a_\varphi]$, $[H \Uparrow^a_\varphi]$ and $[\varphi \rightsquigarrow a]$ is axiomatized by the axiomatization of Theorem 1 as well as the axioms and the rules of Propositions 3, 4, and 6.*

First Application: Tell Action An underlying idea of tell-action is that agent b *privately* tells φ to agent a, that is, the other agents than a will not notice this action. As a result, only agent a will change her belief by φ but the other agents than a will not change their beliefs. After the action, agent a will update her belief not only by the statement φ but also by the signed statement $\mathsf{Sign}(b,\varphi)$. Now we define:

$$[\mathsf{Tell}(b,a,\varphi)]\psi := [\mathsf{Sign}(b,\varphi) \rightsquigarrow a]\psi.$$

Then, we can recover all recursion axioms in [8] by Proposition 6. Especially, we obtain the following.

Proposition 7 (Successful Telling [8]). $[\mathsf{Tell}(b,a,\varphi)]\mathsf{Bel}\big(a, \mathsf{Sign}(b,\varphi)\big)$ *is valid in all si-models.*

This proposition is the essential aspect of tell-action. That is, after agent b tells to agent a information φ, agent a believes that agent b signs φ.

Second Application: Careful Policy In [8], Lorini et al. introduced several policies, as *meta-logical* principles, in order to decide which pieces of information an agent should believe. A common and rational policy is called a *careful policy*. An idea of this policy is to accept, as beliefs, the statements which are universally signed by a group of agents who are equally reliable. Firstly, we define $\mathsf{Sign}(\mathsf{C}^a_i,\varphi)$ which stands for 'all agents who are in the set C^a_i sign statement φ' as follows:

$$\mathsf{Sign}(\mathsf{C}^a_i,\varphi) := \bigwedge\nolimits_{c \in \mathsf{C}^a_i} \big(\mathsf{Sign}(c,\varphi)\big).$$

We also introduce the following abbreviation, whose reading is 'a believes that φ is universally signed by a group of agents who are equally reliable':

$$\mathsf{UniSign}(\varphi,a) := \bigvee\nolimits_{i \leq M} \Big(\mathsf{Bel}\big(a, \mathsf{Sign}(\mathsf{C}^a_i,\varphi)\big) \wedge \mathsf{Bel}\big(a, \bigwedge\nolimits_{1 \leq j \leq i-1} \neg\mathsf{Sign}(\mathsf{C}^a_j,\neg\varphi)\big)\Big),$$

where M is the maximum natural number of $\{i \leq \#G \mid \mathsf{C}_i^a \neq \emptyset\}$. Then, Lorini et al. [8]'s definition of careful policy is introduced as the following implication:

$$\mathsf{UniSign}(\varphi, a) \to \mathsf{Bel}(a, \varphi).$$

However, Lorini et al. did not discuss how we can handle the idea of careful policy in terms of dynamic operators, while they used the policy as a meta-logical principle. With the help of our private announcement operator $[\varphi \rightsquigarrow a]$, we now define the careful policy as a dynamic operator as follows:

$$[\mathsf{Careful}(a, \varphi)]\psi := \mathsf{UniSign}(\varphi, a) \to [\varphi \rightsquigarrow a]\psi,$$

where we may read $[\mathsf{Careful}(a, \varphi)]\psi$ as 'after agent a aggregates information about φ by the careful policy, ψ holds.' By Proposition 6, we obtain the following.

Proposition 8. *The following are valid in all si-models.*

(i) $[\mathsf{Careful}(a, p)]\mathsf{Bel}(a, p)$.
(ii) $[\mathsf{Careful}(a, \mathsf{Sign}(b, \varphi))]\mathsf{Bel}(a, \mathsf{Sign}(b, \varphi))$.

The first item of this proposition says that after agent a aggregates information about p by the careful policy, agent a now believes p. However, we cannot generalize the first item to an arbitrary formula φ, while the second item of this proposition still holds.

4 Dynamic Logical Analysis of Target Legal Case

In order to analyze reliability change from the judge's perspective, we will only focus on the Civil Court. We will not consider the inquiry stage because there is no change of reliability. The Appeal Court and the Supreme Court are also excluded because they only adopted the result of the Civil Court. Furthermore, we will simplify the target legal case by removing agent f_1 in order to avoid an unnecessary complication (this is not an essential point for our analysis).

In the Civil Court, the set G of agents is $\{po, v, f_2, mo, j\}$, where po, v, f_2, mo are agents of four witnesses, and j is a judge of the Civil Court. For the statement involving the legal case, we consider only one propositional letter p whose reading is 'd is the offender' that provides information who is the offender. We assume at first that all witnesses are equally reliable for j as follows:

$$\mathfrak{M}, w \models \mathsf{Bel}(j,\ v \approx_j f_2 \approx_j mo \approx_j po).$$

In the trial, the witness v told a piece of information which is different from the inquiry stage to j. The first action is $T_1 := \mathsf{Tell}\big(v, j, \neg p\big)$. Then, po was called to be a witness and told the received information in the inquiry stage to j that can be represented by the following tell-actions.

$$T_2 := \mathsf{Tell}\big(po, j, \mathsf{Sign}(v, p)\big),\ T_3 := \mathsf{Tell}\big(po, j, \mathsf{Sign}(f_2, \neg p)\big),\ T_4 := \mathsf{Tell}\big(po, j, \mathsf{Sign}(mo, p)\big)$$

After that, j will believe the following information by Proposition 7.

$$\mathfrak{M}, w \models [T_1][T_2][T_3][T_4]\mathsf{Bel}\left(\begin{array}{c} j,\ \mathsf{Sign}\big(v, \neg p\big) \wedge \\ \mathsf{Sign}\big(po, \mathsf{Sign}(v, p) \wedge \mathsf{Sign}(f_2, \neg p) \wedge \mathsf{Sign}(mo, p)\big) \end{array}\right)$$

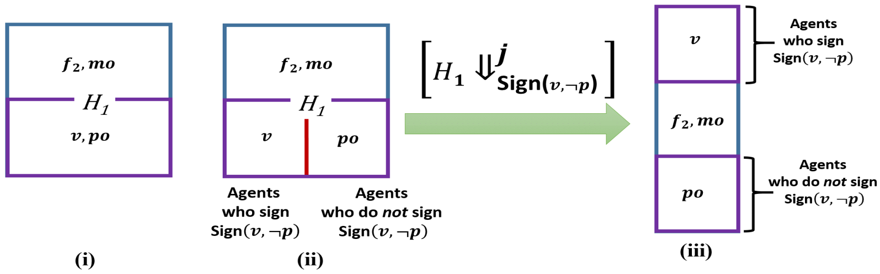

Fig. 2. Downgrading by $[H_1 \Downarrow^{j}_{\mathsf{Sign}(v,\neg p)}]$

Based on these pieces of information alone, j cannot decide which pieces of information should believe. This is firstly because (P1) if j considers the reliability of information sources, j cannot distinguish the reliability ordering between all witnesses because they are equally reliable. Moreover, (P2) there is contradicting information about p in the signed information from all witnesses. So, j cannot decide which signed information should be in j's belief, i.e., p or $\neg p$. We use the following two ideas: (i) reliability change, and (ii) aggregation policy, to resolve the above problems (P1) and (P2).

(i) Reliability change: The downgrade and upgrade operators of Section 3 are applied in order to simulate the effect of reliability change of the judge in the Civil Court.[4] This allows us to solve the above problem (P1). We also note that, if we apply a framework based on [8], a reliability relation between agents is *fixed*, i.e., the reliability relation between agents cannot be changed.

(ii) Aggregation policy: The reformulation of the careful policy (in Section 3) is employed in order to allow the judge of the Civil Court to decide which pieces of the received signed information should believe.

Now let us apply our two ideas to dissolve the judge's difficulty in deciding which pieces of information she should believe. In what follows, we assume that j is the judge in the Civil Court, and define $\mathfrak{M}'$ by the updated model of $\mathfrak{M}$ after the tell-actions T_1–T_4.

From the tell-actions T_1–T_4, there is conflicting information about p. That is, v told statement $\neg p$ (by T_1), while po told signed statement p by v (by T_2). So, j now believes both $\mathsf{Sign}(v, \neg p)$ and $\mathsf{Sign}(po, \mathsf{Sign}(v, p))$. Since the signature operator $\mathsf{Sign}(a, \cdot)$ can be positively introspective, note that $\mathsf{Sign}(v, \neg p)$ implies $\mathsf{Sign}(v, \mathsf{Sign}(v, \neg p))$. From Section 2, we may regard that j believes that the signed information of v in the Civil Court is less reliable than that in the inquiry stage. This means that $\mathsf{Sign}(v, \neg p)$ is not reliable information for j, and so, we regard that j downgrades all agents between po and v who sign the statement $\mathsf{Sign}(v, \neg p)$ by $[H_1 \Downarrow^{j}_{\mathsf{Sign}(v,\neg p)}]$, where we define $H_1 = \{v, po\}$ is a set of agents of witnesses in the Civil Court (see Fig. 2(i)). Let us see a process of the downgrade step by step (Fig. 2). When we consider the agents who sign the statement

[4] In this work, we will not analyze how an agent decides to change the reliability ordering between the other agents.

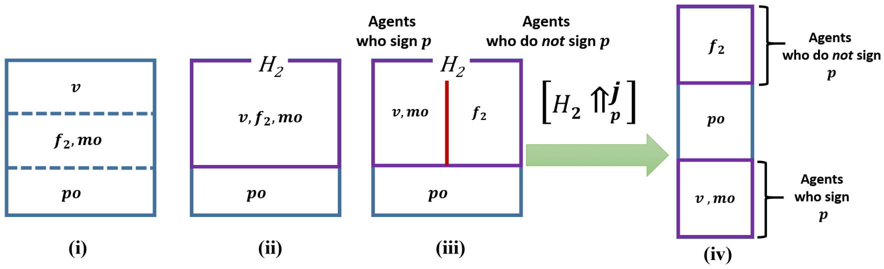

Fig. 3. Upgrading by $[H_2 \Uparrow^j_p]$

$\mathsf{Sign}(v, \neg p)$, H_1 is divided into two equal vertical parts, i.e., $\mathsf{Sign}\big(x, \mathsf{Sign}(v, \neg p)\big)$ and $\neg\mathsf{Sign}\big(x, \mathsf{Sign}(v, \neg p)\big)$ as Fig. 2(ii). Next, j downgrades all agents in H_1 who sign $\mathsf{Sign}(v, \neg p)$ (recall that $\mathsf{Sign}(v, \mathsf{Sign}(v, \neg p))$ holds), and the result can be shown as Fig. 2(iii). That is, the agent v becomes less reliable than all the other agents. Note that the agents who are in the same part are equally reliable. Thus, j changes her belief about the reliability ordering as follows:

$$\mathfrak{M}', w \models [H_1 \Downarrow^j_{\mathsf{Sign}(v,\neg p)}]\mathsf{Bel}(j, po <_j f_2 \approx_j mo <_j v).$$

Since po now becomes the most reliable agent according to j, j can accept the signed statements by po by our careful policy as follows:

$$\mathfrak{M}', w \models [H_1 \Downarrow^j_{\mathsf{Sign}(v,\neg p)}][\mathsf{Careful}\big(j, \mathsf{Sign}(v,p) \wedge \mathsf{Sign}(f_2, \neg p) \wedge \mathsf{Sign}(mo, p)\big)] \\ \mathsf{Bel}\left(j,\ \ \mathsf{Sign}(v,p) \wedge \mathsf{Sign}(f_2, \neg p) \wedge \mathsf{Sign}(mo, p)\right),$$

where we also note that the assumption of the careful policy holds, i.e., $\mathfrak{M}', w \models [H_1 \Downarrow^j_{\mathsf{Sign}(v,\neg p)}]\mathsf{UniSign}(\mathsf{Sign}(v,p) \wedge \mathsf{Sign}(f_2, \neg p) \wedge \mathsf{Sign}(mo, p), j)$ holds. Let us denote $\mathfrak{M}''$ by the updated model of $\mathfrak{M}'$ after the above downgrading and the careful policy.

Since j believes that the signed information of v in the Civil Court is less reliable than that in the inquiry stage, we can regard that j believes that the signed information p of v in the inquiry stage is more reliable. Thus, j upgrades all agents who sign the statement p by $[H_2 \Uparrow^j_p]$, where H_2 is defined by $\{v, f_2, mo\}$ as Fig. 3(ii) (because j focuses on the inquiry stage). Fig. 3 (i) is the initial reliability ordering for j before the upgrading. When we consider the statement p, H_2 is divided into two equal vertical parts, i.e., $\mathsf{Sign}(x,p)$ and $\neg\mathsf{Sign}(x,p)$ as Fig. 3 (iii). By $[H_2 \Uparrow^j_p]$, agents v and mo who sign the statement p are upgraded to be more reliable than all the other agents as Fig. 3(iv). Consequently, j changes her reliability ordering between all witnesses as follows:

$$\mathfrak{M}'', w \models [H_2 \Uparrow^j_p]\mathsf{Bel}(\ j,\ v \approx_j mo <_j po <_j f_2).$$

Since now mo and v become most reliable agents according to j, j now successfully aggregates information p by the careful policy again and will believe that d is the offender (p) as follows:

$$\mathfrak{M}'', w \models [H_2 \Uparrow^j_p][\mathsf{Careful}(j,p)]\ \mathsf{Bel}(j,p).$$

Let us denote $\mathfrak{M}'''$ by the updated model of $\mathfrak{M}''$ after $[H_2 \Uparrow^j_p]$ and $[\mathsf{Careful}(j,p)]$. Therefore, $\mathfrak{M}''', w \models\ \mathsf{Bel}(j,p)$.

5 Conclusion

This work has proposed logical analysis for formalizing reliability change of an agent. We introduced two dynamic operators: downgrading $[H \Downarrow^a_\varphi]$ and upgrading $[H \Uparrow^a_\varphi]$. The first operator downgrades all agents in H who sign φ, while the second operator upgrades them. Based on these operators, we have formalized an example of a legal case from Thailand. In the trials, the judge first believed that all witnesses are equally reliable. Then, the judge changed her belief about the reliability ordering between witnesses. We can successfully analyze this process by downgrading and upgrading the reliability of the witnesses. Moreover, we reformulated the careful policy [8], which allows an agent to decide which signed information should believe, in terms of dynamic operators, i.e., $[\mathsf{Careful}(a, \varphi)]$. Our contribution is to formalize the change of the reliability ordering between the other agents depending on an agent's perspective.

In this work, we only capture an effect of reliability change on belief change, i.e., when a judge changed her reliability ordering between some witnesses, she may change her beliefs about information from those witnesses. On the other hand, belief change may affect reliability change. In this sense, our work just supposes that the judge changes her reliability based on her belief change, but does not analyze how belief change affects reliability change. Therefore, we plan to formalize an effect of belief change on reliability change by applying the notion of preference upgrade in [10]. Furthermore, this work only formalizes the reliability of agents, but does not consider the reliability of statements. That is, this work assumes that when agent a received a statement φ from agent b, agent a has already decided if the statement φ is reliable or not. If agent a considers that the statement φ is not reliable, then she believes that agent b who gives the statement φ will be unreliable. However, we can analyze such reliability change of statements by employing a preference modality based on [10].

References

1. Prakken, H., Sartor, G.: The role of logic in computational models of legal argument - a critical survey. In: Computational Logic: Logic Programming and Beyond. Volume 2408 of Lecture Notes in Computer Science. (2001) 342–381
2. Bench-Capon, T.J.M., Prakken, H.: Introducing the logic and law corner. J. Log. Comput. **18**(1) (2008) 1–12
3. Grossi, D., Rotolo, A.: Logic in the law: A concise overview. Logic and Philosophy Today, Studies in Logic **30** (2011) 251–274
4. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Springer (2008)
5. van Benthem, J.: Dynamic logic for belief revision. Journal of Applied Non-Classic Logics **14**(2) (2004) 129–155
6. Liau, C.J.: Belief, information acquisition, and trust in multi-agent systems–a modal logic formulation. Artificial Intelligence **149**(1) (2003) 31–60
7. Perrussel, L., Lorini, E., Thévenin, J.: From signed information to belief in multi-agent systems. In: Proceedings of The Multi-Agent Logics, Languages, and Organisations Federated Workshops (MALLOW 2010), Lyon, France, August 30 - September 2, 2010. (2010)

8. Lorini, E., Perrussel, L., Thévenin, J.: A modal framework for relating belief and signed information. In: Computational Logic in Multi-Agent Systems - 12th International Workshop, CLIMA XII, Barcelona, Spain, July 17-18, 2011. Proceedings. (2011) 58–73
9. Baltag, A., van Ditmarsch, H.P., Moss, L.S.: Epistemic logic and information update. In Adriaans, P., van Benthem, J., eds.: Handbook on the Philosophy of Information. Elsevier Science Publishers (2008) 361–456
10. van Benthem, J., Liu, F.: Dynamic logic of preference upgrade. Journal of Applied Non-Classical Logics **17**(2) (2007) 157–182
11. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, Cambridge (2001)

A Omitted Proofs

A.1 Proof of Theorem 1

Let us write our axiomatization by $\mathbf{BS}_{\leqslant}$. We show that any unprovable formula φ in $\mathbf{BS}_{\leqslant}$ is falsified in some *si*-model and we basically follow the standard techniques, e.g. found in [11]. Let φ be an unprovable formula in $\mathbf{BS}_{\leqslant}$. We define the canonical model $\mathfrak{M}$ where φ is falsified at some point of $\mathfrak{M}$. We say that a set Γ of formulas is $\mathbf{BS}_{\leqslant}$-*consistent* (for short, *consistent*) if $\bigwedge \Gamma'$ is unprovable in $\mathbf{BS}_{\leqslant}$, for all finite subsets Γ' of Γ, and that Γ is *maximally consistent* if Γ is consistent and $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$ for all formulas φ. Note that ψ is unprovable in $\mathbf{BS}_{\leqslant}$ iff $\neg\psi$ is $\mathbf{BS}_{\leqslant}$-consistent, for any formula ψ. We define the canonical model $\mathfrak{M} = (W, (R_a)_{a\in G}, (S_a)_{a\in G}, (\preccurlyeq_a)_{a\in G}, V)$, for $\mathbf{BS}_{\leqslant}$ by:

- W is the set of all maximal consistent sets;
- $\Gamma R_a \Delta$ iff ($\mathsf{Bel}(a,\psi) \in \Gamma$ implies $\psi \in \Delta$) for all ψ;
- $\Gamma S_a \Delta$ iff ($\mathsf{Sign}(a,\psi) \in \Gamma$ implies $\psi \in \Delta$) for all ψ;
- $b \preccurlyeq_a^{\Gamma} c$ iff $b \leqslant_a c \in \Gamma$;
- $\Gamma \in V(p)$ iff $p \in \Gamma$.

Then, we can show the following equivalence (Truth Lemma [11, Lemma 4.21]): $\mathfrak{M}, \Gamma \models \psi$ iff $\psi \in \Gamma$ for all formulas ψ and $\Gamma \in W$. Given any unprovable formula φ in $\mathbf{BS}_{\leqslant}$, we can find a maximal consistent set Δ such that $\neg\varphi \in \Gamma$. Then, by the equivalence above, φ is falsified at Δ of the canonical model $\mathfrak{M}$ for $\mathbf{BS}_{\leqslant}$, where we can assure that $\mathfrak{M}$ is our intended *si*-model by axioms of Proposition 1. $\square$

A.2 Proof of Proposition 3

We only show that the following four formulas are valid:

$$
\begin{array}{llll}
\text{(A1)} & [H \Downarrow^a_\varphi](b \leqslant_a c) \leftrightarrow & b \leqslant_a c & (b, c \in G \setminus H) \\
\text{(A2)} & [H \Downarrow^a_\varphi](b \leqslant_a c) \leftrightarrow & \big(\mathsf{Sign}(b,\varphi) \wedge \mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c)\big) \vee & \\
& & \big(\neg\mathsf{Sign}(b,\varphi) \wedge \neg\mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c)\big) \vee & \\
& & \big(\neg\mathsf{Sign}(b,\varphi) \wedge \mathsf{Sign}(c,\varphi)\big) & (b, c \in H) \\
\text{(A3)} & [H \Downarrow^a_\varphi](b \leqslant_a c) \leftrightarrow & \mathsf{Sign}(c,\varphi) \vee \big(\neg\mathsf{Sign}(c,\varphi) \wedge (b \leqslant_a c)\big) & (c \in H, b \in G \setminus H) \\
\text{(A4)} & [H \Downarrow^a_\varphi](b \leqslant_a c) \leftrightarrow & \neg\mathsf{Sign}(b,\varphi) \wedge (b \leqslant_a c) & (b \in H, c \in G \setminus H)
\end{array}
$$

Fix any model $\mathfrak{M} = (W, (R_a)_{a\in G}, (S_a)_{a\in G}, (\preccurlyeq_d)_{d\in G}, V)$ and any state $u \in W$. Recall the rewritten $\preccurlyeq'_a$ from the definition of $\mathfrak{M}^{H\Downarrow^a_\varphi}$, i.e., we define $b \preccurlyeq'^u_a c$ iff

(i) $\big(b, c \in H$ and $\mathfrak{M}, u \models \mathsf{Sign}(b,\varphi) \wedge \mathsf{Sign}(c,\varphi)$ and $b \preccurlyeq^u_a c\big)$ or

(ii) $\big(b, c \in (G \setminus H) \cup \{x \in H \mid \mathfrak{M}, u \models \neg\mathsf{Sign}(x,\varphi)\}$ and $b \preccurlyeq^u_a c\big)$ or

(iii) $\big(b \in (G \setminus H) \cup \{x \in H \mid \mathfrak{M}, u \models \neg\mathsf{Sign}(x,\varphi)\}$ and $c \in H$ and $\mathfrak{M}, u \models \mathsf{Sign}(c,\varphi)\big)$

This is equivalent to:

(i) $\big(b, c \in H$ and $\mathfrak{M}, u \models \mathsf{Sign}(b,\varphi) \wedge \mathsf{Sign}(c,\varphi)$ and $b \preccurlyeq^u_a c\big)$ or

(ii1) $\big(b, c \in (G \setminus H)$ and $b \preccurlyeq^u_a c\big)$ or

(ii2) $\big(b, c \in H$ and $\mathfrak{M}, u \models \neg\mathsf{Sign}(b, \varphi)$ and $\mathfrak{M}, u \models \neg\mathsf{Sign}(c, \varphi)$ and $b \preccurlyeq^u_a c\big)$ or
(ii3) $\big(b \in (G \setminus H)$ and $c \in H$ and $\mathfrak{M}, u \models \neg\mathsf{Sign}(c, \varphi)$ and $b \preccurlyeq^u_a c\big)$ or
(ii4) $\big(c \in (G \setminus H)$ and $b \in H$ and $\mathfrak{M}, u \models \neg\mathsf{Sign}(b, \varphi)$ and $b \preccurlyeq^u_a c\big)$ or
(iii1) $\big(b \in (G \setminus H)$ and $c \in H$ and $\mathfrak{M}, u \models \mathsf{Sign}(c, \varphi)\big)$ or
(iii2) $\big(b, c \in H$ and $\mathfrak{M}, u \models \neg\mathsf{Sign}(b, \varphi)$ and $\mathfrak{M}, u \models \mathsf{Sign}(c, \varphi)\big)$

where we note that (ii) is divided into four further cases and (iii) is divided into two further cases. Then we can equivalently rewrite the above seven cases into the following form:

(a1) if $b, c \in G \setminus H$: $b \preccurlyeq^u_a c$ by (ii1)
(a2) if $b, c \in H$: we have either
- $\mathfrak{M}, u \models \mathsf{Sign}(b, \varphi) \wedge \mathsf{Sign}(c, \varphi)$ and $b \preccurlyeq^u_a c$ (by (i)) or
- $\mathfrak{M}, u \models \neg\mathsf{Sign}(b, \varphi)$ and $\mathfrak{M}, u \models \neg\mathsf{Sign}(c, \varphi)$ and $b \preccurlyeq^u_a c$ (by (ii2)) or
- $\mathfrak{M}, u \models \neg\mathsf{Sign}(b, \varphi)$ and $\mathfrak{M}, u \models \mathsf{Sign}(c, \varphi)$ (by (iii2))

(a3) if $c \in H$ and $b \in G \setminus H$: we have either
- $\mathfrak{M}, u \models \mathsf{Sign}(c, \varphi)$(by (iii1)) or
- $\mathfrak{M}, u \models \neg\mathsf{Sign}(c, \varphi)$ and $b \preccurlyeq^u_a c$ (by (ii3))

(a4) if $b \in H$ and $c \in G \setminus H$: $\mathfrak{M}, u \models \neg\mathsf{Sign}(b, \varphi)$ and $b \preccurlyeq^u_a c$ by (ii4)

Then, it is easy to see that we can show that the axiom (Ai) is true in u of $\mathfrak{M}$ in the case (ai) for any $i \in \{1, 2, 3, 4\}$.

A.3 Proof of Proposition 5

The proof of this proposition is similar to the proof of [8, Theorem 2]. First, $\preccurlyeq'_c$ is easily seen to be total pre-ordering. Second, we show that S'_c is still serial, Euclidean and transitive. Recall that

$$(w, e)S'_c(v, f) \text{ iff } wS_cv \text{ and } (e, f) \in U_c.$$

Since both S_c and U_c are both serial, Euclidean and transitive, so is S'_c.

Third, we show that R'_c is still Euclidean and transitive. We concentrate on Euclidean, because the proof of transitivity is also similar. Recall that

$$(w, e)R'_c(v, f) \text{ iff } wR_cv \text{ and } (e, f) \in D_c \text{ and } \mathfrak{M}, v \models \text{pre}(f).$$

To prove that R'_c is Euclidean, fix any (w, e), (v, f), and (u, g) such that $(w, e)R_c(v, f)$ and $(w, e)R_c(u, g)$. By definition, we obtain:

- wR_cv and $(e, f) \in D_c$ and $\mathfrak{M}, v \models \text{pre}(f)$ and
- wR_cu and $(e, g) \in D_c$ and $\mathfrak{M}, u \models \text{pre}(g)$.

To show $(v, f)R'_c(u, g)$, it suffices to show:

- vR_cu and $(f, g) \in D_c$ and $\mathfrak{M}, u \models \text{pre}(g)$.

The third conjunct has already obtained by assumption. vR_cu is shown from wR_cv and wR_cu, since R_c is Euclidean. Similarly, $(f, g) \in D_c$ is also shown from $(e, f) \in D_c$ and $(e, g) \in D_c$, since D_c is Euclidean (for any $c \in G$, see Definition 4).

A.4 Proof of Proposition 6

The most non-trivial part is to show that $[\varphi \rightsquigarrow a]\mathsf{Bel}(c,\psi) \leftrightarrow \mathsf{Bel}(c,\psi)$ $(a \neq c)$ is valid on all *si*-models. Let us fix any model $\mathfrak{M}$ and any state w of $\mathfrak{M}$. We suffice to show

$$\mathfrak{M}, w \models [\varphi \rightsquigarrow a]\mathsf{Bel}(c,\psi) \text{ iff } \mathfrak{M}, w \models \mathsf{Bel}(c,\psi).$$

The right-hand-side is equivalent to:

$$\forall v\,(wR_c v \quad \Rightarrow \quad \mathfrak{M}, v \models \psi) \tag{1.1}$$

The left-hand-side is equivalent to:

$$\begin{aligned}
&\mathfrak{M}^{\varphi \rightsquigarrow a}, (w, !_\varphi) \models \mathsf{Bel}(c,\psi)\\
&\text{iff } \forall (v,f)\,\big((w,!_\varphi)R'_c(v,f) \Rightarrow \mathfrak{M}^{\varphi \rightsquigarrow a}, (v,f) \models \psi\big)\\
&\text{iff } \forall (v,f)\,((wR_c v \text{ and } (!_\varphi, f) \in D_c \text{ and } \mathfrak{M}, v \models \mathrm{pre}(f)) \Rightarrow \mathfrak{M}^{\varphi \rightsquigarrow a}, (v,f) \models \psi)\\
&\text{iff } \forall v\,((wR_c v \text{ and } (!_\varphi, \top) \in D_c \text{ and } \mathfrak{M}, v \models \mathrm{pre}(\top)) \Rightarrow \mathfrak{M}^{\varphi \rightsquigarrow a}, (v,\top) \models \psi) \quad \text{by Definition 5}\\
&\text{iff } \forall v\,((wR_c v \text{ and } (!_\varphi, \top) \in D_c \text{ and } \mathfrak{M}, v \models \top) \Rightarrow \mathfrak{M}^{\varphi \rightsquigarrow a}, (v,\top) \models \psi) \quad \text{by Definition 4}\\
&\text{iff } \forall v\,(wR_c v \Rightarrow \mathfrak{M}^{\varphi \rightsquigarrow a}, (v,\top) \models \psi)\,. \quad \text{by Definition 4}
\end{aligned} \tag{1.2}$$

Therefore, it suffices to establish the equivalences between (1.1) and (1.2). This is reduced to show the following equivalence:

$$\mathfrak{M}, v \models \psi \text{ iff } \mathfrak{M}^{\varphi \rightsquigarrow a}, (v,\top) \models \psi.$$

Let $\mathfrak{M}^{\varphi \rightsquigarrow a}_{(v,\top)}$ be the submodel generated by the point $(v,\top)$ (here the notion of generated submodel are understood in the standard sense of [11]). Then, we suffice to show:

$$\mathfrak{M}, v \models \psi \text{ iff } \mathfrak{M}^{\varphi \rightsquigarrow a}_{(v,\top)}, (v,\top) \models \psi.$$

But, since Definition 4 and Definition 5 imply that $\mathfrak{M}$ and $\mathfrak{M}^{\varphi \rightsquigarrow a}_{(v,\top)}$ are isomorphic (by the mapping sending v of $\mathfrak{M}$ to $(v,\top)$ of $\mathfrak{M}^{\varphi \rightsquigarrow a}_{(v,\top)}$), we can easily obtain the desired equivalence just above. Note that the proof of the validity of $[\varphi \rightsquigarrow a]\mathsf{Sign}(c,\psi) \leftrightarrow \mathsf{Sign}(c,\psi)$ on all *si*-models is similar.

A.5 Proof of Proposition 7

By Proposition 6, we can proceed as follows:

$$\begin{aligned}
[\mathsf{Tell}(b,a,\varphi)]\mathsf{Bel}\big(a, \mathsf{Sign}(b,\varphi)\big) &\leftrightarrow \mathsf{Bel}\big(a, \mathsf{Sign}(b,\varphi) \rightarrow [\mathsf{Tell}(b,a,\varphi)]\mathsf{Sign}(b,\varphi)\big)\\
&\leftrightarrow \mathsf{Bel}\big(a, \mathsf{Sign}(b,\varphi) \rightarrow \mathsf{Sign}(b,\varphi)\big),
\end{aligned}$$

and the last formula is clearly valid on all *si*-models.

A.6 Proof of Proposition 8

For (i), we can proceed by Proposition 6 as follows:

$$\begin{aligned}[\mathsf{Careful}(a,p)]\mathsf{Bel}(a,p) &\leftrightarrow (\mathsf{UniSign}(p,a) \to [p \rightsquigarrow a]\mathsf{Bel}(a,p)) \quad \text{by definition}\\ &\leftrightarrow \mathsf{UniSign}(p,a) \to \mathsf{Bel}(a,p \to [p \rightsquigarrow a]p)\\ &\leftrightarrow \mathsf{UniSign}(p,a) \to \mathsf{Bel}(a,p \to p),\end{aligned}$$

where the last formula is clearly valid on all *si*-models. Next, we show (ii) with the help of Proposition 6 as follows:

$$\begin{aligned}&[\mathsf{Careful}(a,\mathsf{Sign}(b,\varphi))]\mathsf{Bel}(a,\mathsf{Sign}(b,\varphi))\\ \leftrightarrow\ &(\mathsf{UniSign}(\mathsf{Sign}(b,\varphi),a) \to [\mathsf{Sign}(b,\varphi) \rightsquigarrow a]\mathsf{Bel}(a,\mathsf{Sign}(b,\varphi))) \quad \text{by definition}\\ \leftrightarrow\ &\mathsf{UniSign}(\mathsf{Sign}(b,\varphi),a) \to \mathsf{Bel}(a,\mathsf{Sign}(b,\varphi) \to [\mathsf{Sign}(b,\varphi) \rightsquigarrow a]\mathsf{Sign}(b,\varphi))\\ \leftrightarrow\ &\mathsf{UniSign}(\mathsf{Sign}(b,\varphi),a) \to \mathsf{Bel}(a,\mathsf{Sign}(b,\varphi) \to \mathsf{Sign}(b,\varphi)),\end{aligned}$$

where the last formula is clearly valid on all *si*-models.

A.7 Proof of Theorem 2

By $\vdash \psi$ (or $\vdash^{+} \psi$), we mean that ψ is a theorem of the axiomatization $\mathbf{BS}_{\leqslant}$ in the previous proof (or, the axiomatization $\mathbf{BS}^{+}_{\leqslant}$ given in the statement of Theorem 2, respectively.) As for the completeness part, we can reduce the completeness of our dynamic extension to the static counterpart (i.e., Theorem 1) as follows. With the help of the axioms of Propositions 3, 4, and 6, we can define a mapping t sending a formula ψ of the expanded syntax (we denote this by $\mathcal{L}^{+}$ below) possibly with three kinds of dynamic operators (i.e., $[H \Downarrow^{a}_{\varphi}]$, $[H \Uparrow^{a}_{\varphi}]$, and $[\varphi \rightsquigarrow a]$) to a formula $t(\psi)$ of the original syntax $\mathcal{L}$. For this aim, we employ *inside-out strategy*, i.e., we start rewriting the *innermost occurrences* of three kinds of dynamic operators. (So, we do not need to consider an axiom for iterated dynamic operators such as $[\varphi \rightsquigarrow a][\psi \rightsquigarrow a]$ or $[\varphi \rightsquigarrow a][H \Uparrow^{a}_{\varphi}]$.) For example, if one of the innermost dynamic operators is $[\varphi \rightsquigarrow a]$, then we cannot find any occurrences of three kinds of dynamic operators. For inside-out strategy, we need to have the following inference rules for dynamic operators:

$$\frac{\psi \leftrightarrow \psi'}{[H \Downarrow^{a}_{\varphi}]\psi \leftrightarrow [H \Downarrow^{a}_{\varphi}]\psi'} \quad \frac{\psi \leftrightarrow \psi'}{[H \Uparrow^{a}_{\varphi}]\psi \leftrightarrow [H \Uparrow^{a}_{\varphi}]\psi'} \quad \frac{\psi \leftrightarrow \psi'}{[\varphi \rightsquigarrow a]\psi \leftrightarrow [\varphi \rightsquigarrow a]\psi'},$$

to assure the replacement of equivalent formulas inside of a formula. But, these rules are derivable from the corresponding necessitation laws and the reduction axioms for the negation and the conjunction in Propositions 3, 4, and 6. Then, for this mapping t, we can show that $\psi \leftrightarrow t(\psi)$ is valid on all *si*-models and $\vdash^{+} \psi \leftrightarrow t(\psi)$. Then, we can proceed as follows. Fix any formula ψ of $\mathcal{L}^{+}$ such that ψ is valid on all *si*-models. By the validity of $\psi \leftrightarrow t(\psi)$ on all *si*-models, we obtain that $t(\psi)$ is valid on all *si*-models. By Theorem 1, $\vdash t(\psi)$, which implies $\vdash^{+} t(\psi)$. Finally, it follows from $\vdash^{+} \psi \leftrightarrow t(\psi)$ that $\vdash^{+} \psi$, as desired. $\square$

B Responses to Reviewers

Firstly, we would like to thank anonymous referees for valuable comments and suggestions, which provide us to improve the quality of this paper. We have carefully read and revised the paper according to the comments as follows.

B.1 Reviewer 1

This paper starts from a target legal case and designs a formalism to analyze the reliability change in such a case. The reliability change is quite interesting and close to legal practice. I recommend this paper to be accepted. The dynamic logic introduced in the paper is based on the work [8] and the update of preference relation in DEL. Reduction axioms are presented for axiomatizing several logics. The target legal case is also analyzed by using the updating mechanism developed in this paper.

Here I show some points which might be helpful for revising the paper. They are listed as follows:

(1) P.1. Line 2. : which information should believe => which information she should believe (‘‘she" is used in the whole paper.)
(2) P.1 Line 8. : when the judge received => when the judge receives
(3) P2. Line -1. : ‘‘as follows" usually ends with ‘‘:", not a dot. Other places are similar.
(4) P4. Line -8. : the set of agents is finite. It is not so clear what is the idea of [8] for using a universal quantifier for agents. Is the whole system in this paper a subsystem of [8] since the universal quantifier is not used? More precisely, is there translation from your system to a fragment of the system of [8]? It seems better to give some further comments on your work and [8].
(5) P5. Below table 1. : ‘‘with our syntax" => ‘‘for our syntax"
(6) P5. Line +6 below table 1. : $\preccurlyeq_a$ is a map for the agent a, but it says ‘‘... for each agent". I suggest to rewrite this sentence.
(7) P6. Line 6. : The concept of rank is not defined here. It makes difficulty for understanding the ranking. In Line 14, it says that the C_1^a differs from [8], but it is unclear what is the true difference.
(8) P6. Theorem 1. : This theorem is stated here without proof. It is better to give a reference or sketch for the proof. The modal logic contains two modalities which do not interact with each other. But there are also axioms for the binary relation $\leq_a$. If we use the canonical model for proving the completeness, how shall we define the canonical model?
(9) P7. Line -6. : In the third case, I think the condition $b \preccurlyeq_a^u c$ is missed.
(10) P8. Line 5. : In the third case, I think the condition $b \preccurlyeq_a^u c$ is missed.

```
(11) P8. Proposition 2. : This proposition is stated without proof. The
     most important part should be the proof of properties of those binary
     relations. Since those properties are universal and the updated
     reliability relation is a subrelation of the original one, we can
     conclude the proposition. It is better to give a sketch of the proof
     here.
(12) P8. Proposition 3. : Again, there is no proof of this proposition.
     It seems to be better to prove the most important cases. Moreover,
     the recursive axioms for iterated dynamic operators are not given
     here. There two dynamic operators here. So there should be four
     recursive axioms for iterated dynamic operators. If they are not
     supplied, Theorem 2 could be problematic. If we want to give a complete
     axiomatization without using recursive axioms for iterated dynamic
     operators, one possible way is to give more rules for those dynamic
     operators. So I suggest to repair this theorem.
(13) P9. Line 12. : ``pre assigns a precondition to each action"
(14) P9. Proposition 5. : Again, no proof. Like the previous comment,
     the iterated axioms are missed. The situation here is more complicated.
     You have three dynamic operators.
(15) P10. Proposition 7. : Does the proposition hold for all formulas
     and agents?
(16) P.11 Line 8. : It says that the item (i) cannot be generalized to
     arbitrary formula. Why? Could you give a short example?
```

From the above comments, we will divide into nine issues as follows.

1. According to items (1), (2), (3), (5), (6) and (13), there are many minor language problems.
   $\Longrightarrow$ Thank you for pointing out this issue. In our new version, we have revised our paper and edited all of language problems.

2.) According to the item (4), the set of agents is finite. It is not so clear what is the idea of [8] for using a universal quantifier for agents. Is the whole system in this paper a subsystem of [8] since the universal quantifier is not used? More precisely, is there translation from your system to a fragment of the system of [8]? It seems better to give some further comments on your work and [8].
   $\Longrightarrow$ Thank you for pointing out this issue. One of the syntactical differences of [8] from ours is that Lorini et al. use the notion of variables for agents to use the quantifier for agents, and Lorini et al. allow the quantification over agents of the modality $\mathsf{Sign}(\cdot, \varphi)$. However, because the set of agents in [8] is still *finite* (not infinite), we realized that the use in [8] of universal quantifier over agents are *redundant*. Actually, when the second author visited Amsterdam in November 2014, he had a chance to discuss this issue with Emiliano Lorini. Consequently, Lorini also agreed with that his use of the quantifier is not necessary for his study in [8]. In this sense, our work can also be regarded as a nice simplification of the previous work, i.e., we remove

an unnecessary syntactic component from the previous work and make its contribution more explicit.

3. According to the item (7), the concept of rank is not defined here. It makes difficulty for understanding the ranking. In Line 14, it says that the C_1^a differs from [8], but it is unclear what is the true difference.
$\Longrightarrow$ For this comment, we have added some explanations about the concept of rank and added the reason why C_1^a differs from [8] in the first paragraph of Page 6 (our additional description is shown as the bold text) as follows:

   "Based on Definition 1 and the idea of [8], **we can rank agents by giving a partition $(\mathsf{C}_i^a)_{i \leq M}$ to G, where M is a natural number representing the maximum rank (such M always exists because G is finite) and we read $c \in \mathsf{C}_i^a$ as 'from agent a's viewpoint, the rank of agent c is i'.** As a result, the agents who are equally reliable are categorized in the same group. C_1^a which stands for 'a group of agents which is the most reliable from a's perspective' can be defined by the following formula:

$$c \in \mathsf{C}_1^a =_{def} \bigwedge\nolimits_{b \in G} (c \leqslant_a b),$$

   where we recall that G is a finite set of agents and $a, b, c \in G$. **Note that we relativize the notion C_i^a to a specified agent a because the notion of reliability ordering $\leqslant_a$ depends on a specified agent a. This is a difference from [8] because [8] did not consider C_i depending on a specified agent.** Then, we can rank the group of agents C_i^a such that $i > 1$ as follows:"

4. According to the item (8), Theorem 1. : This theorem is stated here without proof. It is better to give a reference or sketch for the proof. The modal logic contains two modalities which do not interact with each other. But there are also axioms for the binary relation $\leq_a$. If we use the canonical model for proving the completeness, how shall we define the canonical model?
$\Longrightarrow$ Thank you for pointing out this issue. We have added the proof of Theorem 1 in Appendix A.1 as follows:

   Let us write our axiomatization by $\mathbf{BS}_{\leqslant}$. We show that any unprovable formula φ in $\mathbf{BS}_{\leqslant}$ is falsified in some *si*-model and we basically follow the standard techniques, e.g. found in [11]. Let φ be an unprovable formula in $\mathbf{BS}_{\leqslant}$. We define the canonical model $\mathfrak{M}$ where φ is falsified at some point of $\mathfrak{M}$. We say that a set Γ of formulas is $\mathbf{BS}_{\leqslant}$-*consistent* (for short, *consistent*) if $\bigwedge \Gamma'$ is unprovable in $\mathbf{BS}_{\leqslant}$, for all finite subsets Γ' of Γ, and that Γ is *maximally consistent* if Γ is consistent and $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$ for all formulas φ. Note that ψ is unprovable in $\mathbf{BS}_{\leqslant}$ iff $\neg\psi$ is $\mathbf{BS}_{\leqslant}$-consistent, for any formula ψ. We define the canonical model $\mathfrak{M} = (W, (R_a)_{a \in G}, (S_a)_{a \in G}, (\preccurlyeq_a)_{a \in G}, V)$, for $\mathbf{BS}_{\leqslant}$ by:

- W is the set of all maximal consistent sets;
- $\Gamma R_a \Delta$ iff $(\mathsf{Bel}(a, \psi) \in \Gamma$ implies $\psi \in \Delta)$ for all ψ;
- $\Gamma S_a \Delta$ iff $(\mathsf{Sign}(a, \psi) \in \Gamma$ implies $\psi \in \Delta)$ for all ψ;
- $b \preccurlyeq_a^{\Gamma} c$ iff $b \leqslant_a c \in \Gamma$;
- $\Gamma \in V(p)$ iff $p \in \Gamma$.

Then, we can show the following equivalence (Truth Lemma [11, Lemma 4.21]): $\mathfrak{M}, \Gamma \models \psi$ iff $\psi \in \Gamma$ for all formulas ψ and $\Gamma \in W$. Given any unprovable formula φ in $\mathbf{BS}_{\leqslant}$, we can find a maximal consistent set Δ such that $\neg\varphi \in \Gamma$. Then, by the equivalence above, φ is falsified at Δ of the canonical model $\mathfrak{M}$ for $\mathbf{BS}_{\leqslant}$, where we can assure that $\mathfrak{M}$ is our intended *si*-model by axioms of Proposition 1. □

5. According to item (9), P7. Line -6. : In the third case, I think the condition $b \preccurlyeq_a^u c$ is missed. According to item (10), P8. Line 5. : In the third case, I think the condition $b \preccurlyeq_a^u c$ is missed.
$\Longrightarrow$ For this issue, we think that the reviewer may misunderstand about the third case in Definition 2 and Definition 3 because we do not have any explanations. Thus, we have added some comments in the footnote of Page 7 as follows:

   "In this case, since there is no relation between agents b and c, $b \preccurlyeq_a^u c$ is omitted."

6. According to items (11), this issue is about the proof of Proposition 2.
$\Longrightarrow$ Thank you for pointing out this issue. First, we note that Definitions 2 and 3 (in Section 3.2) for upgrading and downgrading rigorously describe the intuitive idea explained in Fig. 1 and the relevant paragraph. We also would like to emphasize that our intuitive description itself explains why we can still keep the properties of total pre-ordering after upgrading or downgrading. This explains why Proposition 2 holds. Since our notion of upgrading or downgrading are similar to the notion of *radical upgrade* studied by van Benthem, we also note that the resulting total pre-ordering is not a subrelation of the original one.

7. According to items (12), and (14), this issue is about the proof of Propositions 3 and 5, respectively.
$\Longrightarrow$ Thank you for pointing out this issue. In order to reflect this issue, we have demonstrated by the proofs of Propositions 3, Propositions 5, and Theorem 2 in Appendixes A.2, A.3 and A.7, respectively.

8. According to the item (15), Proposition 7. : Does the proposition hold for all formulas and agents?
$\Longrightarrow$ Thank you for pointing out this issue. From Proposition 7, this proposition holds for all formulas and agents. See also the proof of Proposition 7 in Appendix A.5.

9. According to the item (16), P.11 Line 8. : It says that the item (i) cannot be generalized to arbitrary formula. Why? Could you give a short example?
   $\Longrightarrow$ Thank you for pointing out this issue. For example, if we define a formula φ by $p \wedge \neg\mathsf{Bel}(a, p)$, then $[\mathsf{Careful}(a, \varphi)]\mathsf{Bel}(a, \varphi)$ cannot hold, since the rewritten equivalent formula (by Proposition 6) becomes

$$\mathsf{UniSign}(\varphi, a) \to \mathsf{Bel}(a, (p \to \mathsf{Bel}(a, p))),$$

   which is not valid in all *si*-models.

B.2 Reviewer 2

```
This paper proposes a logical framework for analyzing reliability change
of an agent, based on the former work [8]. Introducing the operations
upgrading and downgrading is the originality in this paper. This paper
is relevant to Juris-Informatics.
   This paper is well-organized. It would be better if the authors first
explain why the former model [8] is difficult to analyze the target
legal case, which will reveal the problem to be solved.
   In paragraph 2 in Section 3: A set of formulas of L is defined with
some modifications. Although I'm not sure how the elimination of the
universal quantifier affects the framework, it simply seems the proposed
framework restricts the set of agents to finite. Is it necessary for
the formalization? If so, the authors should mention the reason of restriction.
Otherwise, meaningless modification from the base model is not preferable.
   Intuitively, figures seem upside-down. The operations upgrading and
downgrading should make agents move to upward and downward, respectively.
```

In order to reflect the above comments, we will split into three issues as follows.

1. It would be better if the authors first explain why the former model [8] is difficult to analyze the target legal case, which will reveal the problem to be solved.
   $\Longrightarrow$ Thank you for your comment. In Section 2 and Section 4, we have regarded that there is reliability change of the judge in the Civil Court. That is, the judge first believed that all witnesses are equally reliable. Then, the judge changed her belief about the reliability ordering between witnesses. In Section 1, we have addressed a limitation of the former model [8] in the second paragraph (shown as the bold text) as follows:

   > "Recently, many studies [1–3] presented the use of logical approaches in the legal systems. Dynamic epistemic logic (DEL) [4, 5] is a logical tool to study reasoning about information change due to communication between agents. This paper focuses on reliability change. There are several works [6–8] that proposed a logical framework for formalizing the reliability. Among of them, Lorini et al. [8] introduced a modal framework for reasoning about signed information. In their framework, the agents can keep track of the information source by using the notion of signed statement. They also considered the notion of reliability over the information sources. **However, they did not deal with the dynamics of the reliability relation of agents.**"

   With the above limitation of [8], we were confronted with the following problem (P1) mentioned in Section 4 (shown as the bold text).

"Based on these pieces of information alone, j cannot decide which pieces of information should believe. This is firstly because **(P1) if j considers the reliability of information sources, j cannot distinguish the reliability ordering between all witnesses because they are equally reliable.** Moreover, (P2) there is contradicting information about p in the signed information from all witnesses. So, j cannot decide which signed information should be in j's belief, i.e., p or $\neg p$. We use the following two ideas: (i) reliability change, and (ii) aggregation policy, to resolve the above problems (P1) and (P2).

(i) Reliability change: The downgrade and upgrade operators of Section 3 are applied in order to simulate the effect of reliability change of the judge in the Civil Court. This allows us to solve the above problem (P1). **We also note that, if we apply a framework based on [8], a reliability relation between agents is *fixed*, i.e., the reliability relation between agents cannot be changed.**

(ii) Aggregation policy: The reformulation of the careful policy (in Section 3) is employed in order to allow the judge of the Civil Court to decide which pieces of the received signed information should believe."

The above description shows that the former model [8] is difficult to analyze the target legal case because it cannot capture reliability change of an agent.

2. In paragraph 2 in Section 3: a set of formulas of $\mathcal{L}$ is defined with some modifications. Although I'm not sure how the elimination of the universal quantifier affects the framework, it simply seems the proposed framework restricts the set of agents to finite. Is it necessary for the formalization? If so, the authors should mention the reason of restriction. Otherwise, meaningless modification from the base model is not preferable.
$\Longrightarrow$ Thank you for pointing out this issue. Firstly, we will describe why we remove the use of the universal quantifier as follows. One of the syntactical differences of [8] from ours is that Lorini et al. use the notion of variables for agents to use the quantifier for agents, and Lorini et al. allow the quantification over agents of the modality $\mathsf{Sign}(\cdot, \varphi)$. However, because the set of agents in [8] is still *finite* (not infinite), we realized that the use in [8] of universal quantifier over agents are *redundant*. Actually, when the second author visited Amsterdam in November 2014, he had a chance to discuss this issue with Emiliano Lorini. Consequently, Lorini also agreed with that his use of the quantifier is not necessary for his study in [8]. For this reason, our work (without the use of the universal quantifier) can be regarded as a nice simplification of the previous work, i.e., we remove an unnecessary syntactic component from the previous work and make its contribution more explicit.

3. Intuitively, figures seem upside-down. The operations upgrading and downgrading should make agents move to upward and downward, respectively.

$\Longrightarrow$ For this comments, we assumed that Fig.1 in Section 3.2 seems upside-down, i.e., the top of the rectangle represents the agents who are less reliable, while the bottom of the rectangle represents the agents who are the most reliable. Therefore, an operation of downgrading makes the agents move to upward (see Fig.1(iii)). On the other hand, an operation of upgrading make the agents move to downward (see Fig.1(iv)).

B.3 Reviewer 3

The paper present a logic mechanism for changing the reliability of agents, when taking into account the statements they make.The work is based on the modal framework introduced by Lorini et al.

While the work provides a way of modeling legal proceedings, it is unclear to me from reading the paper what the benefit of doing this is. In case of the Thai example, other decisions by the judge could be represented as well. Overall, the motivation for the work is not clear to me.

The paper is technically sound.

The paper suffers in readability because of lack of intuition (section 3.3) and issues with the English language (missing articles, verb and subjects not corresponding). Comments (suggested additions in capital):

- abstract: lacks motivation
- p1: which information should IT believe
- p1: The process of the trials: Trial proceedings
- p1: and so on: delete
- p1: keep track of THE information
- p2: example of A legal case (in) FROM Thailand
- p2: in THE Trang ...
- p2: bleeding at the lung: rephrase
- p2 and p3: The more details: More details
- p3: is the offender and intendED to kill v FOR (by) the
- p3: scrapper belongs: belongs -> is
- p3: d uses it for attacking that would be lethal: d uses it in a possibly lethal attack
- p3: such dangerous resulting in death: possibly lethal
- p4: imprisoned by death: ??
- p4: on (the) previous work
- p4: our example (in) FROM A logical
- p4: formalise THE belief change
- p6: who sign(ED) the statement ϕ, ψ: verb missing
- p7: fig 1 (iii): why is there a gap? while there is no gap in (iv) between the two groups
- p7: c_1 and c_2: add , after c_2
- p8: this section reviews THE tell-action
- p9: aggregates policy: policies
- p9: called THE careful policy
- p10: that is, (the) ALL other agents (than a)
- p10: As a result, ONLY agent a will ... (but the other ...)
- p11: adopted: adapted?
- p12: reliable agent ACCORDING to
- p13: agents ACCORDING to j
- p13: an example of A legal case
- p14: how belief change affect THE reliability ...

From the above comments, we will separate into the following issues.

1. While the work provides a way of modeling legal proceedings, it is unclear to me from reading the paper what the benefit of doing this is. In case of the Thai example, other decisions by the judge could be represented as well. Overall, the motivation for the work is not clear to me.
$\Longrightarrow$ For this comment, we have some explanations about the motivation of our work in the first paragraph of Section 1 as follows.

   > "In agent communication, an agent needs some criteria to decide which information she should believe. A common criterion is to consider the reliability of an information source. If the agent considers that a source of received information is reliable, she would accept and might believe the received information. On the other hand, the agent may reject the received information if she considers that the source is not reliable. Legal proceedings are a typical example of agent communication that the reliability has a strong influence on a judge's decision in a court. Thus, the judge needs the reliability, i.e., when a judge receives a piece of information from a witness, the judge should consider if the witness is reliable or not. In addition, when the judge receives new information, she might change her belief about the reliability of the witness. This paper aims to investigate an effect of reliability change of the judge in legal judgment."

   The above description shows that our goal is to investigate an effect of reliability change of the judge in legal judgment. That is, our work demonstrates the simulation of the judge's decision in order to make the reader easy to understand the legal case. This is the benefit of our work.

2. The paper suffers in readability because of lack of intuition (section 3.3).
$\Longrightarrow$ Thank you for pointing out this issue. We have added some description of this issue in the first paragraph of Section 3.3 (our additional description is shown as the bold text) as follows.

   > "**This section introduces a new dynamic operator for private announcement $[\varphi \rightsquigarrow a]$ (whose reading is "after a private announcement of φ to agent a"). One of the merits of this operator is that we do not specify information of sender of message φ, while agent a is a recipient of the message. This means that we may use this operator also for self-decision of agent a, i.e., the sender and the recipient are the same.** This section demonstrates that $[\varphi \rightsquigarrow a]$ can capture both (i) the *tell*-action $[\mathsf{Tell}(b, a, \varphi)]$ from [8]: 'agent b tells to agent a that a certain statement φ is true' and (ii) one of aggregation policies from [8] called the *careful policy*: 'accept, as beliefs, the statements which are universally signed by a group of agents who are equally reliable'. **We note that Lorini et al. [8] did not propose a logical treatment**

from dynamic epistemic viewpoints for any of aggregation policies. Moreover, we note that the sender and the recipient are regarded as the same to capture the careful policy by our new operator $[\varphi \rightsquigarrow a]$**.**"

3. The paper suffers in readability because of issues with the English language (missing articles, verb and subjects not corresponding).
$\Longrightarrow$ Thank you for pointing out this issue. In our new version, we have revised our paper and edited all of English language problems. However, there are some comments that we did not edit corresponding to the reviewer as follows:
   - p1: which information should IT believe. In this issue, since the another reviewer also comments about this issue and he/she suggests that we should write as "which information SHE should believe", we think that this sentence seems to be correct. So, we would like to rewrite the sentence as "which information SHE should believe".
   - p2: in THE Trang. In this issue, since Trang a province in Thailand, we do not use THE.
   - p4: formalise THE belief change. In this issue, since we assume that belief change can be regarded as general ideas, we do not use THE when we refer general ideas.
   - p7: c_1 and c_2: add , after c_2. Since this issue is unclear to us, we will rewrite this sentence as "agents b_1 and b_2 which are equally reliable are more reliable than agents c_1 and c_2 which are equally reliable from agent a's perspective.".
   - p10: that is, (the) ALL other agents (than a). In this issue, we did not removed 'than a' because we want to emphasize this point in order to clearly understand.
   - p10: As a result, ONLY agent a will ... (but the other ...). In this issue, we added a word 'ONLY', but did not removed 'but the other ...' because we want to emphasize this point in order to clearly understand.
   - p11: adopted: adapted? In this issue, we think that a word 'adopted' is correctly used for representing that the Appeal Court and the Supreme Court only adopted the result from the Civil Court, i.e., there is no need to adjust or alter.
   - p14: how belief change affect THE reliability. In this issue, since we assume that reliability change can be regarded as general ideas, we do not use THE when we refer general ideas.

4. From the first item of comments, this paper lacks motivation in the abstract.
$\Longrightarrow$ Thank you for pointing out this issue. We have rewritten the abstract and added the motivation at the beginning of the abstract (our additional description is shown as the bold text) as follows.

   "**A consideration of the reliability plays a significant role in agent communication. An agent can change her belief about**

> **the reliability ordering between the other agents with respect to new incoming information.** In order to analyze reliability change of an agent, this paper proposes a logical formalization with two dynamic operators, i.e., downgrade and upgrade operators. The downgrade operator allows an agent to downgrade some specified agents to be less reliable corresponding to the degree of reliability, while the upgrade operator allows an agent to upgrade them to be more reliable. Furthermore, we demonstrate our formalization by a legal case from Thailand."

5. From p7, fig 1 (iii): why is there a gap? while there is no gap in (iv) between the two groups.
   $\Longrightarrow$ From Section 3.2, we have written the description of downgrading and upgrading as follows:

   > "Before giving a detailed semantics, let us demonstrate the effect of $[H \Downarrow^a_\varphi]$ and $[H \Uparrow^a_\varphi]$ by figures. Firstly, we assume that a rectangle G of Fig. 1(i) represents a fixed finite set of agents. Secondly, we will select a specified set of agents in order to change their reliability ordering that can be represented by a rectangle H, and we assume that $b_1 \approx_a b_2 <_a c_1 \approx_a c_2$ hold, i.e., agents b_1 and b_2 which are equally reliable are more reliable than agents c_1 and c_2 which are equally reliable from agent a's perspective. In this sense, b_1, b_2, c_1 and c_2 are situated as in Fig. 1(i). Then, if we focus on the agents who sign the statement φ, H is divided into two equal vertical parts, i.e., $\mathsf{Sign}(x, \varphi)$ and $\neg\mathsf{Sign}(x, \varphi)$ as in Fig. 1(ii), namely by the set $\{x \in H \mid \mathfrak{M}, w \models \mathsf{Sign}(x, \varphi)\}$ and the set $\{x \in H \mid \mathfrak{M}, w \models \neg\mathsf{Sign}(x, \varphi)\}$. Next, if the agent a downgrades all the agents signing the statement φ, we downgrade all of them less reliable than the other agents as in Fig.1(iii). On the other hand, if the agent a upgrades all the agents signing the statement φ, we upgrade all of them more reliable than the other agents as in Fig.1(iv)."

   From the above description, we can regard that G is a fixed finite set of agents and H is a specified set of agents in order to change their reliability ordering. In Fig.1(i) and Fig.1(ii), a gap can be represented by $G \setminus H$ that stands for a set of agents which are in G and are not in H. For Fig.1(iii), the agents signing the statement φ are downgraded to be less reliable than the other agents, i.e., they are moved to be the top of the rectangle. On the other hand, in Fig.1(iv), the agents signing the statement φ are upgraded to be more reliable than the other agents, i.e., they are moved to be the bottom of the rectangle.