Price of Privacy

Pavel Naumov

Vassar College, Poughkeepsie, New York, USA

Jia Tao

The College of New Jersey, Ewing, New Jersey, USA

Abstract

The article proposes a logical framework for reasoning about agents' ability to protect their privacy by hiding certain information from a privacy intruder. It is assumed that the knowledge of the intruder is derived from the observation of pieces of evidence and that there is a cost associated with the elimination of the evidence. The logical framework contains a modal operator labeled by a group of agents and a total budget available to this group. The key contribution of this work is the proposed incorporation of the cost factor into privacy protection reasoning within the standard modal logic framework. The main technical result are the soundness and completeness theorems for the introduced logical system with respect to a formally defined semantics.

1. Introduction

Privacy and Costs. The cost associated with maintaining privacy is a topic of public [11, 5] and scholarly [12, 17, 14, 18] discussions. There are at least two aspects in our daily life where costs are explicitly or implicitly associated with people's privacy.

On one hand, catering to the increasing desire to protect afore-disclosed personal information, several companies[1] offer services of removing references to such information from public sites, search engines, and commercial databases for a fee. Some of them[2] offer an additional service of disseminat-

Email addresses: pnaumov@vassar.edu (Pavel Naumov), taoj@tcnj.edu (Jia Tao)

[1] reputation.com, abinedeleteme.com

[2] reputation.com/reputationdefender

Preprint submitted to Journal of Applied Logic *November 18, 2016*

ing positive information about a person on the Internet so that the negative information is harder to find. Similarly, before mobile phones became popular, close to one third of American households paid monthly fee to have their number unlisted [1]. Nowadays, the unlisted phone service is still being offered by the phone companies and for significantly higher monthly fees than before [6].

On the other hand, consumers often reveal their private information unintentionally to companies in exchange for a small discount by using mail-in rebates, coupons, or store discount cards. Such information may later be analyzed for marketing purposes. For example, the second-largest US discount retailer Target developed an approach to identify pregnant women by tracking their shopping patterns of seemingly not-baby-related items such as scent-free soap and extra-big bags of cotton balls [2]. In these cases, consumers usually have an option not to use the promotional discount, and thus to pay a bit more, but to avoid the disclosure of their private information. In practice, this option of preserving privacy for an additional cost is rarely used by consumers, possibly due to the lack of awareness.

The price that people have to pay for protecting their privacy may differ from one individual to another. For example, European "right to be forgotten" law [9] makes it essentially free for individuals to remove certain information from online search engines. At the same time, the removal of similar information in the United States might be impossible or achievable only by paying significant legal fees.

Modal Language. In this article we introduce a logical system for reasoning about costs of protecting privacy by hiding some knowledge from a given privacy intruder. We assume that the information is being hidden from a single fixed privacy intruder that often will be referred to as just "the intruder". In the conclusion we talk about possible extensions of our logical systems to handle multiple privacy intruders.

To specify such a logical system, one could consider modality $\mathbf{H}^c_a\varphi$ with meaning "at cost c agent a can *hide* φ from the intruder". Such a modality, however, does not satisfy the standard Necessitation rule from modal logic:

$$\frac{\varphi}{\mathbf{H}^c_a\varphi}$$

for any value of c. To observe this, assume that formula φ is a propositional tautology. For example, let φ be of the form $\psi \vee \neg\psi$. Being a propositional tautology formula $\psi \vee \neg\psi$ is universally true. At the same time, for each non-negative value c, formula $\mathbf{H}_a^c(\psi \vee \neg\psi)$ is not true because no matter what actions with total cost c are taken by agent a to hide $\psi \vee \neg\psi$ from the privacy intruder, it is still known to the intruder by the virtue of being a propositional tautology.

To solve this issue, in this article we use modality $\Box_a^c\varphi$ that stands for "at cost c agent a *cannot* hide φ from the intruder", which is the negation of the "hiding" modality: $\Box_a^c\varphi \equiv \neg\mathbf{H}_a^c\varphi$. This modality does satisfy the Necessitation axiom

$$\frac{\varphi}{\Box_a^c\varphi}$$

because if φ is universally true, then, as we have just discussed above, its knowledge cannot be hidden by the agent a from anyone at any cost.

As usual in modal logic, one can also define dual modality $\Diamond_a^c\varphi$ as $\neg\Box_a^c\neg\varphi$. Under our semantics statement $\Diamond_a^c\varphi$ is interpreted as "at cost c agent a can leave the intruder under an impression that φ could be true". Note that

$$\mathbf{H}_a^c\varphi \equiv \neg\Box_a^c\varphi \equiv \Diamond_a^c\neg\varphi. \tag{1}$$

In other words, hiding φ means creating an impression that $\neg\varphi$ could be true. The formal semantics of these modalities will be given in Definition 6.

Second Order Privacy. Statement $\mathbf{H}_a^c\varphi$ says that agent a can hide information φ from the intruder at cost c. Some agents, especially corporate entities, treat their business costs as a tightly guarded secret. Such agents might be interested not only in hiding φ, but also in hiding how much it costs to them to hide φ. The "second order" hiding (the hiding of the costs of hiding) also has a cost associated with it. Generally speaking, the latter cost is unrelated to the former one. The fact that at cost d the agent a can hide that at costs c she can hide φ could be expressed in our language through nested modalities as $\mathbf{H}_a^d\mathbf{H}_a^c\varphi$.

Suppose that an agent a hires a privacy protection company (agent b) to hide information φ at cost c. The privacy protection company might be interested to keep secret the price it charges agent a. Doing so might be an additional business expense d for the privacy company. Thus, $\mathbf{H}_b^d\mathbf{H}_a^c\varphi$.

Note that nested modality $\mathbf{H}_b^d\mathbf{H}_a^c\varphi$ does *not* represent a joint effort by the two agents to hide information φ. Instead, the two nested modalities refer to hiding two different facts by the agents. We next extend our modal language to capture the "joint effort".

Agents Cooperation. If A is any finite set of agents, then by $\mathbf{H}^c_A\varphi$ we denote that the agents in set A, working together, can hide information φ from the intruder at a total cost c. For the reason discussed earlier, we have chosen modality $\Box^c_A\varphi \equiv \neg\mathbf{H}^c_A\varphi$ to be the primitive construction in our language. Statement $\Box^c_A\varphi$ means that agents in set A, working together, *cannot* hide information φ from the intruder at a total cost c. The single-agent notation $\Box^c_a\varphi$ that was used before can now be formally interpreted as $\Box^c_{\{a\}}\varphi$.

Principles of Privacy Protection. The main result of this article is a sound and complete logical system that describes properties of the cost of protecting privacy by a group of agents through hiding some information from the intruder. Even though our logical system essentially consists of the axioms of epistemic logic S5 for distributed knowledge [4] with an addition of cost superscript, the meaning of our axioms is quite different from the meaning of S5 axioms.

Our version of the *Truth axiom*:

$$\Box^c_A\varphi \to \varphi$$

states that if a group of agents A cannot hide fact φ from the intruder for some cost c, then statement φ is true. The same principle can be rephrased in terms of modality $\mathbf{H}$ as formula $\neg\varphi \to \mathbf{H}^c_a\varphi$. Recall that hiding of φ means creating an impression that $\neg\varphi$ could be true. Since formula $\neg\varphi$ is actually true by the assumption of our rephrased principle, creating the impression that $\neg\varphi$ might be true does not take agent a any effort and does not cost her anything. Thus, $\mathbf{H}^c_a\varphi$. We will make this argument more formal when we prove the soundness of the Truth axiom with respect to a formally defined semantics in Lemma 3.

The *Positive Introspection axiom*

$$\Box^{c+d}_B\varphi \to \Box^c_A\Box^d_B\varphi,$$

where $A \subseteq B$, states that if a larger group B cannot hide φ at cost $c+d$, then any subgroup $A \subseteq B$ cannot hide from the intruder, at cost c, the fact that group B cannot hide φ at cost d. This principle can be rephrased in the terms of the hiding modality as $\mathbf{H}^c_A\neg\mathbf{H}^d_B\varphi \to \mathbf{H}^{c+d}_B\varphi$. To understand why this principle is valid, recall that hiding φ means creating an impression that $\neg\varphi$ could be true. Thus, per the assumption of our rephrased principle, agents in set A at cost c can create an impression that $\mathbf{H}^d_B\varphi$ could be true. In other words, agents in set A at cost c can create an impression that agents in set B at cost d can create an impression that φ is false. Working together, agents in sets A and B at combined cost $c+d$ can create an impression that φ is false. That is, $\mathbf{H}^{c+d}_{A\cup B}\varphi$. Therefore, $\mathbf{H}^{c+d}_B\varphi$ due to the assumption $A \subseteq B$. We will make this argument more formal when we prove the soundness of the Positive Introspection axiom with respect to a formally defined semantics in Lemma 4.

The *Negative Introspection axiom* is true in our setting in the following form:

$$\neg\Box^c_B\varphi \to \Box^d_A\neg\Box^{c+d}_B\varphi,$$

where $A \subseteq B$. The axiom states that if a larger group B *can* hide φ at cost c, then any subgroup $A \subseteq B$ *cannot* hide from the intruder, at cost d, the fact that group B *can* hide φ at cost $c+d$. The axiom could be rephrased in terms of hiding modality as

$$\mathbf{H}^d_A\mathbf{H}^{c+d}_B\varphi \to \neg\mathbf{H}^c_B\varphi.$$

The claim of this axiom is perhaps counterintuitive and cannot be easily justified without giving formal semantics of information hiding. We prove the soundness of this axiom with respect to the formally defined semantics in Lemma 5.

Finally, the *Distributivity axiom*,

$$\Box^c_A(\varphi \to \psi) \to (\Box^c_A\varphi \to \Box^c_A\psi),$$

states that if a group of agents A can hide neither $\varphi \to \psi$ nor φ from the intruder at cost c, then it cannot hide ψ at the same cost c. Informally, validity of this principle is self-evident. We prove the soundness of this axiom with respect to the formally defined semantics in Lemma 6.

Related Literature. The hiding of knowledge from the intruder is closely related to forgetting the knowledge. Different ways of forgetting from the epistemic point of view have been discussed by van Ditmarsch et al. [16]. They proposed a sound and complete logical system of propositional variable forgetting. Although their logical system has a more sophisticated semantics of forgetting than our semantics of evidence elimination, their system does not take into account the cost. Zhang and Zhou [19] proposed an alternative semantics of forgetting and gave four semantic postulates completely characterizing their notion of forgetting. However, due to the semantical nature of these postulates, they do not form a logical system. Zhang and Zhou's approach also does not consider the cost associated with forgetting.

There are very few papers that combine the cost of actions and logic. Endriss et al. [3] investigated taxation schemes in Boolean games. In our work on budget constrained knowledge [7], we interpreted modality $\Box^c_a\varphi$ as "at cost c agent a can learn that fact φ is true". In spite of the semantic similarity between the budget constrained knowledge and the knowledge hiding, these two modalities have very different properties: the budget constrained modality does not satisfy the Negative Introspection axiom and it has a different form of Distributivity axiom: $\Box^c_a(\varphi \to \psi) \to (\Box^d_a\varphi \to \Box^{c+d}_a\psi)$.

This article is also related to our article on the logic of confidence [8], where we interpreted modality $\Box^c_a\varphi$ as "agent a knows that statement φ is true assuming that she makes measurements with precision $\pm c$". Unlike the current article that considers modalities labeled by sets of agents, logic of confidence only considers modalities labeled by single agents. However, if one restricts the logical system in the current article to single agent labels, then the resulting logical system, although with a totally different semantics, would become very similar to the logic of confidence. The only difference would be that the logic of confidence contains "Zero Confidence" axiom: $\varphi \to \Box^0_a\varphi$ which is not sound under the "knowledge hiding" semantics of the current article. This is because of our assumption that some pieces of evidence might be eliminated at zero cost. We have taken this approach for the sake of generality.

The article is organized as follows. In the next section we introduce the syntax and formal semantics of our logical system. Than we list axioms of this system. In the two sections that follow, we prove the soundness and the completeness of our logical systems, respectively. The last section concludes.

2. Syntax and Semantics

Throughout the article we assume a fixed set of atomic propositions $\mathcal{P}$ and a fixed set of agent variables $\mathcal{A}$. We start by defining the language Φ of our formal system.

Definition 1. *Let set Φ be the minimal set of formulas such that*

1. *$p \in \Phi$ for each atomic proposition $p \in \mathcal{P}$,*
2. *$\varphi \to \psi \in \Phi$ for each $\varphi, \psi \in \Phi$,*
3. *$\neg\varphi \in \Phi$ for each $\varphi \in \Phi$,*
4. *$\Box^c_A\varphi \in \Phi$ for each non-negative real number c, each nonempty finite subset $A \subseteq \mathcal{A}$, and each $\varphi \in \Phi$.*

Note that we only allow non-negative real values of c in the modality $\Box^c_A$.

We next introduce a formal semantics in which groups of agents can hide some information from a privacy intruder. We assume that the intruder derives its knowledge by observing certain pieces of evidence. For example, a discount store can conclude that a woman is pregnant by observing that she buys either scent-free soap or an extra-big bag of cotton balls[3]. From an epistemic logic point of view, each piece of evidence can be treated as an indistinguishability relation between epistemic worlds. That is, the presence of scent-free soap in a woman's shopping record can be used to distinguish epistemic worlds where the shopping record contains this item from the worlds where it does not contain the item.

For example, consider two epistemic worlds w and u such that Mary buys scent-free soap in both worlds, Jane buys scent-free soap in neither of these worlds, and Kathy buys scent-free soap in world w, but not in world u. In this setting,

$$w \sim_{\text{“Mary buys scent-free soap.”}} u,$$

$$w \sim_{\text{“Jane buys scent-free soap.”}} u,$$

$$w \nsim_{\text{“Kathy buys scent-free soap.”}} u.$$

Representing pieces of evidence as equivalence relations on epistemic worlds was introduced in our work on budget-constrained knowledge [7]. It is, however, not the only way to define formal semantics of evidence. Van Benthem and Pacuit use neighborhood models [15]. Neighborhood models of evidence are geared towards treating beliefs. For example, they assume that evidence does not necessarily support the current world. Such semantics would not fit into the scope of our paper since we use word “evidence” as a synonym for justifications of true facts.

[3] This is a significantly simplified version of the model used by Target to predict pregnancy [2].

Following the assumption that the intruder derives its knowledge by observing certain pieces of evidence, we formally interpret the hiding of knowledge as the elimination of these pieces of evidence. We further assume that there is a cost of elimination of an evidence to each agent. For example, if scent-free soap is 20 cents off when using a store discount card, a woman can hide her pregnancy by not using the card and, thus, paying 20 cents more but eliminating the evidence of her pregnancy. Generally speaking, not every agent is able to eliminate every evidence. If an agent cannot erase a piece of evidence, then we say that the cost of eliminating this piece to the agent is infinity. Thus, our semantics allows infinite cost of evidence elimination although our syntax only allows non-negative real values of c in the modality $\Box^c_A$.

Definition 2. *A tuple $\langle \mathcal{W}, \mathcal{E}, \{\sim_e\}_{e\in\mathcal{E}}, \{\|\cdot\|_a\}_{a\in\mathcal{A}}, \pi\rangle$ is called a Kripke model if*

1. *$\mathcal{W}$ is an arbitrary set (of epistemic worlds),*
2. *$\mathcal{E}$ is an arbitrary set (of evidences),*
3. *$\sim_e$ is an (indistinguishability) equivalence relation on the set $\mathcal{W}$ associated with the piece of evidence $e \in \mathcal{E}$,*
4. *$\|\cdot\|_a$ is a (cost) function from set $\mathcal{E}$ into set $\{r \in \mathbb{R} \mid r \geq 0\} \cup \{+\infty\}$, for each agent $a \in \mathcal{A}$,*
5. *π is a function from atomic propositions into subsets of $\mathcal{W}$.*

As another example, consider a scenario where Mary and Jane, both pregnant, are standing in a cashier line. Mary has scent-free soap in her shopping cart and Jane has an extra-big bag of cotton balls. The soap is 20 cents off and cotton balls are 30 cents off when a store discount card is presented. In this setting, each woman can eliminate the evidence of her own pregnancy by not using the discount card. However, she cannot eliminate the evidence of the other woman's pregnancy:

$$\begin{aligned}
&\|\text{“Mary buys scent-free soap.”}\|_{Mary} = 0.20,\\
&\|\text{“Mary buys scent-free soap.”}\|_{Jane} = +\infty,\\
&\|\text{“Jane buys a big bag of cotton balls.”}\|_{Jane} = 0.30,\\
&\|\text{“Jane buys a big bag of cotton balls.”}\|_{Mary} = +\infty.
\end{aligned}$$

Definition 3. $\|e\|_A = \min_{a\in A} \|e\|_a$, *for every $e \in \mathcal{E}$ and every $A \subseteq \mathcal{A}$.*

For example, in the situation described above,

$$\begin{aligned}&\|\text{“Mary buys scent-free soap.”}\|_{\{Mary,Jane\}} = 0.20,\\&\|\text{“Jane buys a big bag of cotton balls.”}\|_{\{Mary,Jane\}} = 0.30.\end{aligned}$$

Definition 4. $\|E\|_A = \sum_{e\in E}\|e\|_A$, *for every* $E \subseteq \mathcal{E}$ *and every* $A \subseteq \mathcal{A}$.

For instance,

$$\begin{aligned}&\|\{\text{“Jane buys a big bag of cotton balls.”},\\&\quad\text{“Mary buys scent-free soap.”}\}\|_{\{Mary,Jane\}} = 0.50.\end{aligned}$$

Next, we prove two technical lemmas that are used later to show the soundness of our logical system.

Lemma 1. $\|E_1\cup E_2\|_A \le \|E_1\|_A+\|E_2\|_A$, *for every finite subsets* $E_1, E_2 \subseteq \mathcal{E}$ *and every finite subset* $A \subseteq \mathcal{A}$.

Proof.

$$\|E_1 \cup E_2\|_A = \sum_{e\in E_1\cup E_2}\|e\|_A \le \sum_{e\in E_1}\|e\|_A + \sum_{e\in E_2}\|e\|_A = \|E_1\|_A + \|E_2\|_A.$$

⊠

Lemma 2. $\|E\|_B \le \|E\|_A$, *for every finite subset* $E \subseteq \mathcal{E}$ *and every two finite subsets* A *and* B *of* $\mathcal{A}$ *such that* $A \subseteq B$.

Proof.

$$\|E\|_B = \sum_{e\in E}\|e\|_B = \sum_{e\in E}\min_{x\in B}\|e\|_x \le \sum_{e\in E}\min_{x\in A}\|e\|_x = \sum_{e\in E}\|e\|_A = \|E\|_A.$$

⊠

For any set of pieces of evidence E, we write $w \sim_E u$ if epistemic worlds w and u cannot be distinguished based on the pieces of evidence in set E. For example, if epistemic worlds w and u differ only by a third woman, Kathy, being pregnant in w and not pregnant in u, then these two epistemic worlds cannot be distinguished by the discount store based on the items in Mary's and Jane's shopping carts: $w \sim_{\{e_1,e_2\}} u$, where

$$\begin{aligned} e_1 &= \text{“Mary buys scent-free soap.”}, \quad \text{and} \\ e_2 &= \text{“Jane buys a big bag of cotton balls.”} \end{aligned}$$

Formally, this relation is defined as follows:

Definition 5. *For any worlds $w, u \in \mathcal{W}$ and any subset $E \subseteq \mathcal{E}$, let $w \sim_E u$ mean that $w \sim_e u$ for each $e \in E$.*

Corollary 1. *Relation $\sim_E$ is an equivalence relation on set $\mathcal{W}$ for each subset $E \subseteq \mathcal{E}$.* $\Box$

Recall from the Introduction that the intended meaning of $w \Vdash \Box^c_A \varphi$ is “group of agents A constrained by a total budget c cannot hide φ from the intruder”. In this article, we formally capture hiding information as removing evidence. Thus, under our formal semantics, $w \Vdash \Box^c_A \varphi$ means that “group of agents A constrained by a total budget c cannot remove enough pieces of evidence to hide φ from the intruder”.

For example, if w is the epistemic world in which Mary is pregnant and is buying both scent-free soap (20 cents off with her shopping card) and an extra-big bag of cotton balls (30 cents off), then paying only extra 40 cents is not sufficient for her to hide her pregnancy from the store:

$$w \Vdash \Box^{0.40}_{\{Mary\}} \text{“Mary is pregnant.”}$$

However, she can hide her pregnancy by not using her shopping card and, thus, paying extra 50 cents:

$$w \Vdash \neg\Box^{0.50}_{\{Mary\}} \text{“Mary is pregnant.”}$$

If u is an epistemic world in which pregnant Mary is buying scent-free soap and Jane, also pregnant, is buying an extra-big bag of cotton balls, then when constrained by 40-cent budget they *cannot* hide the fact that at least one of them is pregnant, but they *can* hide the fact that they are both pregnant if one of them does not use her discount card:

$$u \Vdash \Box^{0.40}_{\{Mary,Jane\}}\text{“One of Mary and Jane is pregnant.”}$$
$$u \Vdash \neg\Box^{0.40}_{\{Mary,Jane\}}\text{“Mary and Jane are both pregnant.”}$$

The first of the above claims is true because no matter which evidence with a cost up to 40 cents (or a set of evidences with a total cost up to 40 cents) is removed, it still can be inferred that at least one of them is pregnant. In other words, after the removal of any set of pieces of evidence with total cost up to 40 cents, at least one of these two women is pregnant in each epistemic world indistinguishable to the store from world u. This is formally captured in item 4 of the definition below.

Definition 6. *For any formula $\varphi \in \Phi$ and any world $w \in \mathcal{W}$ of a Kripke model $\langle \mathcal{W}, \mathcal{E}, \{\sim_e\}_{e\in\mathcal{E}}, \{\|\cdot\|_a\}_{a\in\mathcal{A}}, \pi\rangle$, let the satisfiability relation $w \Vdash \varphi$ be defined as follows:*

1. *$w \Vdash p$ if $w \in \pi(p)$,*
2. *$w \Vdash \neg\psi$ if $w \nVdash \psi$,*
3. *$w \Vdash \psi \to \chi$ if $w \nVdash \psi$ or $w \Vdash \chi$,*
4. *$w \Vdash \Box^c_A\psi$ if $u \Vdash \psi$, for each finite $E \subseteq \mathcal{E}$ such that $\|E\|_A \leq c$ and each $u \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E} u$.*

3. Axioms

Our logical system, in addition to the propositional tautologies in language Φ, contains the following axioms for each sets A and B such that $A \subseteq B$:

1. Truth: $\Box^c_A\varphi \to \varphi$,
2. Positive Introspection: $\Box^{c+d}_B\varphi \to \Box^c_A\Box^d_B\varphi$,
3. Negative Introspection: $\neg\Box^c_B\varphi \to \Box^d_A\neg\Box^{c+d}_B\varphi$,
4. Distributivity: $\Box^c_A(\varphi \to \psi) \to (\Box^c_A\varphi \to \Box^c_A\psi)$.

We write $\vdash \varphi$ if formula φ is provable from the propositional tautologies and the above axioms using Modus Ponens and Necessitation inference rules:

$$\frac{\varphi, \quad \varphi \to \psi}{\psi} \qquad\qquad \frac{\varphi}{\Box^c_A\varphi}\ .$$

We write $X \vdash \varphi$ if formula φ is provable from propositional tautologies, the above axioms, and the additional set of axioms X using only Modus Ponens inference rule.

Next, we give two examples of proofs in our logical system.

Proposition 1. *$\vdash \Box^c_A\varphi \to \Box^d_A\varphi$, for each $c \geq d \geq 0$, each subset $A \subseteq \mathcal{A}$, and each $\varphi \in \Phi$.*

Proof. By the Positive Introspection axiom, $\vdash \Box^c_A\varphi \to \Box^{c-d}_A\Box^d_A\varphi$. By the Truth axiom, $\vdash \Box^{c-d}_A\Box^d_A\varphi \to \Box^d_A\varphi$. Then, from the two statements above using propositional logic we can conclude that $\vdash \Box^c_A\varphi \to \Box^d_A\varphi$. ☒

Proposition 2. *$\vdash \Box^c_B\varphi \to \Box^c_A\varphi$, for each $c \geq 0$, each $\varphi \in \Phi$, and each pair of subsets A and B of $\mathcal{A}$ such that $A \subseteq B$.*

Proof. By the Truth axiom, $\vdash \Box^0_B\varphi \to \varphi$. Hence, by the Necessitation rule, $\vdash \Box^c_A(\Box^0_B\varphi \to \varphi)$. Thus, by the Distributivity axiom and the Modus Ponens inference rule,

$$\vdash \Box^c_A\Box^0_B\varphi \to \Box^c_A\varphi. \tag{2}$$

At the same time, by the Positive Introspection axiom, $\vdash \Box^c_B\varphi \to \Box^c_A\Box^0_B\varphi$. Therefore, $\vdash \Box^c_B\varphi \to \Box^c_A\varphi$ using statement (2) and propositional logic. ☒

4. Soundness

In this section we establish the soundness of our logical system with respect to the semantics given in Definition 6. The soundness of propositional tautologies and of Modus Ponens inference rule is straightforward. Below we prove the soundness of each of the remaining axioms and of Necessitation inference rule as separate lemmas. We assume that (i) w is an arbitrary epistemic world of a Kripke model $\langle \mathcal{W}, \mathcal{E}, \{\sim_e\}_{e\in\mathcal{E}}, \{\|\cdot\|_a\}_{a\in\mathcal{A}}, \pi\rangle$, (ii) $\varphi, \psi \in \Phi$, (iii) $A, B \subseteq \mathcal{A}$, and (iv) $c, d \geq 0$.

Lemma 3 (Truth). *If $w \Vdash \Box^c_A\varphi$, then $w \Vdash \varphi$.*

Proof. By Definition 6, assumption $w \Vdash \Box^c_A\varphi$ implies that $u \Vdash \varphi$ for each finite $E \subseteq \mathcal{E}$ such that $\|E\|_A \leq c$ and each $u \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E} u$.

Consider $E = \varnothing$. By Corollary 1, relation $\sim_{\mathcal{E}\setminus\varnothing}$ is an equivalence relation on set $\mathcal{W}$. Thus, $w \sim_{\mathcal{E}\setminus\varnothing} w$. Also note that

$$\|E\|_A = \sum_{e\in\varnothing} \|e\|_A = 0 \leq c.$$

Therefore, $w \Vdash \varphi$. $\boxtimes$

Lemma 4 (Positive Introspection). *If $w \Vdash \Box^{c+d}_B\varphi$ and $A \subseteq B$, then $w \Vdash \Box^c_A\Box^d_B\varphi$.*

Proof. Consider any $E_1 \subseteq \mathcal{E}$ such that $\|E_1\|_A \leq c$ and any $u \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E_1} u$. It suffices to show that $u \Vdash \Box^d_B\varphi$. To prove this, consider any $E_2 \subseteq \mathcal{E}$ such that $\|E_2\|_B \leq d$ and any $v \in \mathcal{W}$ such that $u \sim_{\mathcal{E}\setminus E_2} v$. We need to show that $v \Vdash \varphi$.

First, note that $w \sim_{\mathcal{E}\setminus E_1} u$ implies $w \sim_{\mathcal{E}\setminus(E_1\cup E_2)} u$. Similarly, $u \sim_{\mathcal{E}\setminus E_2} v$ implies $u \sim_{\mathcal{E}\setminus(E_1\cup E_2)} v$. By Corollary 1, relation $\sim_{\mathcal{E}\setminus(E_1\cup E_2)}$ is transitive. Thus, $w \sim_{\mathcal{E}\setminus(E_1\cup E_2)} v$.

Second, by Lemma 1 and Lemma 2,

$$\|E_1 \cup E_2\|_B \leq \|E_1\|_B + \|E_2\|_B \leq \|E_1\|_A + \|E_2\|_B \leq c + d.$$

By Definition 6, assumption $w \Vdash \Box^{c+d}_B\varphi$ implies that $w' \Vdash \varphi$ for each $E \subseteq \mathcal{E}$ such that $\|E\|_B \leq c + d$ and each $w' \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E} w'$. Consider, in particular, $E' = E_1 \cup E_2$ and $w' = v$. Then, $v \Vdash \varphi$. $\boxtimes$

Lemma 5 (Negative Introspection). *If $w \nVdash \Box^c_B\varphi$ and $A \subseteq B$, then $w \Vdash \Box^d_A\neg\Box^{c+d}_B\varphi$.*

Proof. By Definition 6, assumption $w \nVdash \Box^c_B\varphi$ implies that there is $E_1 \subseteq \mathcal{E}$ such that $\|E_1\|_B \leq c$ and there is $u \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E_1} u$ and $u \nVdash \varphi$.

Consider any $E_2 \subseteq \mathcal{E}$ such that $\|E_2\|_A \leq d$ and any $v \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E_2} v$. By Definition 6, it suffices to prove that $v \nVdash \Box^{c+d}_B\varphi$. Note $w \sim_{\mathcal{E}\setminus E_1} u$ implies that $w \sim_{\mathcal{E}\setminus(E_1\cup E_2)} u$. Similarly, $w \sim_{\mathcal{E}\setminus E_2} v$ implies that $w \sim_{\mathcal{E}\setminus(E_1\cup E_2)} v$. Hence, $v \sim_{\mathcal{E}\setminus(E_1\cup E_2)} u$ by Corollary 1. At the same time, $u \nVdash \varphi$ and, by Lemma 1 and Lemma 2,

$$\|E_1 \cup E_2\|_B \leq \|E_1\|_B + \|E_2\|_B \leq \|E_1\|_B + \|E_2\|_A \leq c + d.$$

Therefore, $v \nVdash \Box^{c+d}_{B}\varphi$ by Definition 6. $\boxtimes$

Lemma 6 (Distributivity). *If $w \Vdash \Box^{c}_{A}(\varphi \to \psi)$ and $w \Vdash \Box^{c}_{A}\varphi$, then $w \Vdash \Box^{c}_{A}\psi$.*

Proof. Consider any finite $E \subseteq \mathcal{E}$ such that $\|E\|_A \leq c$ and any $u \in W$ such that $w \sim_{\mathcal{E}\setminus E} u$. By Definition 6, it suffices to prove that $u \Vdash \psi$. Indeed, assumptions $w \Vdash \Box^{c}_{A}\varphi$ and $w \Vdash \Box^{c}_{A}(\varphi \to \psi)$, by Definition 6, imply that $u \Vdash \varphi$ and $u \Vdash \varphi \to \psi$. Therefore, again by Definition 6, $u \Vdash \psi$. $\boxtimes$

Lemma 7 (Necessitation). *If $w' \Vdash \varphi$ for each epistemic world w' of each Kripke model, then $w \Vdash \Box^{c}_{A}\psi$.*

Proof. Consider any $E \subseteq \mathcal{E}$ such that $\|E\|_A \leq c$ and any $u \in W$ such that $w \sim_{\mathcal{E}\setminus E} u$. By Definition 6, it suffices to show that $u \Vdash \varphi$, which is true due to the assumption of the lemma. $\boxtimes$

5. Completeness

In the rest of this article, we prove the completeness theorem for our logical system that is stated later as Theorem 1. The proof of the completeness is based on the "unravelling" technique [13].

For an arbitrary maximal consistent subset s_0 of set Φ, we define a Kripke model

$$\mathcal{K}(s_0) = \langle \mathcal{W}, \mathcal{E}, \{\sim_e\}_{e\in\mathcal{E}}, \{\|\cdot\|\}_{a\in\mathcal{A}}, \pi\rangle$$

that will be referred to as the *canonical model.*

We start with a formal definition of set $\mathcal{W}$, followed by an informal discussion of the intuition behind this definition.

Definition 7. *Let the set of epistemic worlds $\mathcal{W}$ be the set of all sequences*

$$\langle s_0, (A_1, d_1), s_1, (A_2, d_2), s_2, \dots, (A_n, d_n), s_n\rangle$$

such that $0 \leq n$ and for each $0 < i \leq n$,

1. *s_i is a maximal consistent subset of Φ,*

2. *the finite nonempty subset of agents* $A_i \subseteq \mathcal{A}$ *and the real number* $d_i \geq 0$ *are such that the following two conditions are satisfied for each* $c \geq 0$, *each* $B \supseteq A_i$, *and each* $\varphi \in \Phi$*:*
   (a) *if* $\Box_B^{c+d_i}\varphi \in s_{i-1}$, *then* $\Box_B^{c}\varphi \in s_i$,
   (b) *if* $\Box_B^{c+d_i}\varphi \in s_i$, *then* $\Box_B^{c}\varphi \in s_{i-1}$.

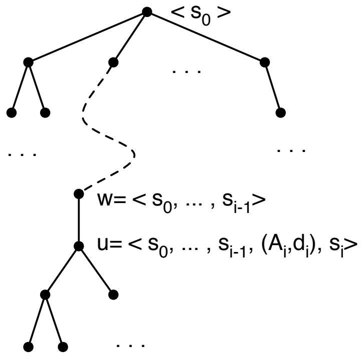
$< s_0 >$
...
...
...
$w = < s_0, \ldots , s_{i-1} >$
$u = < s_0, \ldots , s_{i-1}, (A_i,d_i), s_i >$
...

Figure 1: Canonical Tree.

Intuitively, the set of epistemic worlds $\mathcal{W}$ can be viewed as a tree, see Figure 1, which we call the *canonical tree*, where world

$$w = \langle s_0, (A_1, d_1), s_1, \ldots, (A_{i-1}, d_{i-1}), s_{i-1}\rangle$$

is a parent node of world

$$u = \langle s_0, (A_1, d_1), s_1, \ldots, (A_{i-1}, d_{i-1}), s_{i-1}, (A_i, d_i), s_i\rangle$$

and world $\langle s_0\rangle$ is the root of the tree. In Definition 9 below, we define the set of pieces of evidence $\mathcal{E}$ for this model in a way that there is a set of pieces of evidences $E_{w,u} \subseteq \mathcal{E}$ such that $\|E_{w,u}\|_{A_i} = d_i$ and $w \sim_{\mathcal{E}\setminus E_{w,u}} u$. In other words, epistemic worlds w and u are not distinguishable if the set of pieces of evidence $E_{w,u}$ is eliminated.

To understand the reason behind condition (a) of item 2 in Definition 7, consider statement $\Box_B^{c+d_i}\varphi \in s_{i-1}$ in the canonical model. We would like it to mean that statement $w \Vdash \Box_B^{c+d_i}\varphi$ is true. Hence, by Definition 6, for each $E \subseteq \mathcal{E}$ such that $\|E\|_B \leq c + d_i$ and each $v \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E} v$, we have $v \Vdash \varphi$. Note that $A_i \subseteq B$, $w \sim_{\mathcal{E}\setminus E_{w,u}} u$, and $\|E_{w,u}\|_{A_i} = d_i$. Thus, for each $E' \subseteq \mathcal{E}$ such that $\|E'\|_B \leq c$ and each $v \in \mathcal{W}$ such that $u \sim_{\mathcal{E}\setminus E'} v$ we have $v \Vdash \varphi$. Therefore, $u \Vdash \Box^c_B\varphi$, which in a canonical model we would like to mean that $\Box^c_B\varphi \in s_i$. Condition (b) of item 2 is defined for the similar reason.

The next corollary directly follows from Definition 7.

Corollary 2. *For any $0 \leq k \leq n$, any $c \geq 0$, any formula $\varphi \in \Phi$, any set $B \subseteq \mathcal{E}$, and any $\langle s_0, (A_1, d_1), s_1, (A_2, d_2), s_2, \dots, (A_n, d_n), s_n\rangle \in \mathcal{W}$, such that $\bigcup_{i=k+1}^n A_i \subseteq B$,*

1. *if $\Box_B^{c+d_{k+1}+\dots+d_n}\varphi \in s_k$, then $\Box^c_B\varphi \in s_n$,*
2. *if $\Box_B^{c+d_{k+1}+\dots+d_n}\varphi \in s_n$, then $\Box^c_B\varphi \in s_k$.*

In the next definition we introduce several technical notations that are used throughout the proof of the completeness.

Definition 8. *If w is an epistemic world*

$$\langle s_0, (A_1, d_1), s_1, (A_2, d_2), s_2, \dots, (A_n, d_n), s_n\rangle,$$

then let

1. *$\sigma(w) = s_n$,*
2. *$\alpha(w) = A_n$, if $n > 0$,*
3. *$\delta(w) = d_n$, if $n > 0$.*

Note that $\alpha(\langle s_0\rangle)$ and $\delta(\langle s_0\rangle)$ are not defined.

Recall that for each pair of epistemic worlds w and u such that node w is the parent of node u in the canonical tree, we intend to have a set of pieces of evidence $E_{w,u}$ such that $\|E_{w,u}\|_{\alpha(u)} = \delta(u)$ and $w \sim_{\mathcal{E}\setminus E_{w,u}} u$. The pieces of evidence in set $E_{w,u}$ are all pairs (u, a) such that $a \in \alpha(u)$. This leads us to the following definition of the set of all pieces of evidence in the canonical model.

Definition 9. *The evidence set $\mathcal{E}$ is $\{(u, a) \mid u \in \mathcal{W} \setminus \{\langle s_0\rangle\}, a \in \alpha(u)\}$.*

We say that world $w_1 \in \mathcal{W}$ is a *prefix* of world $w_2 \in \mathcal{W}$ (denoted by $w_1 \preceq w_2$) if for some $m \leq n$, world w_1 is equal to $\langle s_0, (A_1, d_1), s_1, \dots, (A_m, d_m), s_m\rangle$ and world w_2 is equal to $\langle s_0, (A_1, d_1), s_1, \dots, (A_m, d_m), s_m, \dots, (A_n, d_n), s_n\rangle$. If $m < n$, then we say that world w_1 is a *proper prefix* of world w_2 and denote it by $w_1 \prec w_2$. In other words, $w_1 \preceq w_2$ means that node w_1 is an ancestor of node w_2 in the canonical tree.

Lemma 8. *For any worlds $u, w_1, w_2 \in \mathcal{W}$, if $u \neq w_2$,*

$$\begin{aligned} w_1 &= \langle s_0, (A_1, d_1), s_1, \dots, (A_{m-1}, d_{m-1}), s_{m-1} \rangle, \text{ and} \\ w_2 &= \langle s_0, (A_1, d_1), s_1, \dots, s_{m-1}, (A_m, d_m), s_m \rangle, \end{aligned}$$

then $u \preceq w_1$ if and only if $u \preceq w_2$.

Proof. Any prefix of sequence w_1 is also a prefix of sequence w_2. The only prefix of sequence w_2 which is not a prefix of sequence w_1 is sequence w_2 itself. Together, these two statements imply the claim of the lemma. ☒

We now define the indistinguishability relations for the canonical model.

Definition 10. *For any worlds $w_1, w_2 \in \mathcal{W}$ and any piece of evidence $(u, a) \in \mathcal{E}$, the indistinguishability relation $w_1 \sim_{(u,a)} w_2$ holds if the following two conditions are either both true or both false: (i) $u \preceq w_1$, (ii) $u \preceq w_2$.*

Lemma 9. *For each $(u, a) \in \mathcal{E}$, relation $\sim_{(u,a)}$ is an equivalence relation on set $\mathcal{W}$.*

Proof. The statement of the lemma follows from Definition 10 and the fact that bi-conditional "either both true or both false" is an equivalence relation in Boolean algebra. ☒

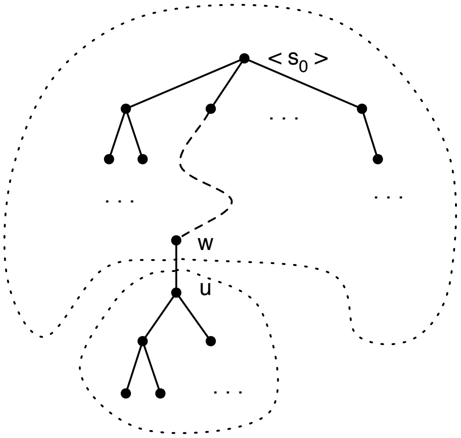

Figure 2: Two equivalence classes of relation $\sim_{(u,a)}$.

Relation $\sim_{(u,a)}$ partitions epistemic worlds into two equivalence classes. One class contains all sequences v for which u is a prefix and the other class contains all other epistemic worlds. In other words, if w is the parent node of u in the canonical tree, as depicted in Figure 2, then the classes are the sets of nodes of the two connected components obtained by removing edge (w, u) from the canonical tree.

Definition 11. *For any two sequences $w_1, w_2 \in \mathcal{W}$, the greatest common prefix $gcp(w_1, w_2)$ is the longest sequence $u \in \mathcal{W}$ such that $u \preceq w_1$ and $u \preceq w_2$.*

In terms of the canonical tree, $gcp(w_1, w_2)$ is the most recent common ancestor of nodes w_1 and w_2. The next lemma characterizes relation $w_1 \sim_{(u,a)} w_2$ in terms of greatest common prefixes.

Lemma 10. *For any $a \in \alpha(v)$, $w_1 \nsim_{(v,a)} w_2$ if and only if $gcp(w_1, w_2) \prec v \preceq w_1$ or $gcp(w_1, w_2) \prec v \preceq w_2$.*

Proof. ($\Rightarrow$) Suppose that $w_1 \nsim_{(v,a)} w_2$. Thus, by Definition 10, without loss of generality, assume that $v \preceq w_1$ and $v \npreceq w_2$. By Definition 11, $gcp(w_1, w_2) \preceq w_1$. So, sequences v and $gcp(w_1, w_2)$ are both prefixes of w_1. There are two possible cases to consider (see Figure 3):

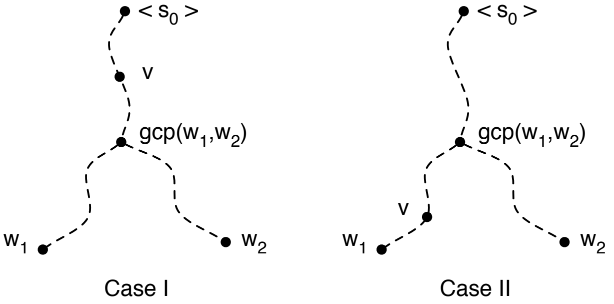

Figure 3: Towards the proof of Lemma 10.

Case I: If $v \preceq gcp(w_1, w_2)$, then, $v \preceq gcp(w_1, w_2) \preceq w_2$, which is a contradiction with our assumption that $v \npreceq w_2$.
Case II: If $gcp(w_1, w_2) \prec v$, then, $gcp(w_1, w_2) \prec v \preceq w_1$.
($\Leftarrow$) Without loss of generality, suppose that $gcp(w_1, w_2) \prec v \preceq w_1$. Thus, sequence v is a prefix of sequence w_1, but not of sequence w_2. Hence, $v \npreceq w_2$.

By Definition 10, statements $v \preceq w_1$ and $v \not\preceq w_2$ imply that $w_1 \nsim_{(v,a)} w_2$. ⊠

As we pointed out earlier, equivalence relation $\sim_{(u,a)}$ partitions epistemic worlds into two equivalence classes (see Figure 2). Informally, it means that evidence (u, a) can be used to distinguish two epistemic worlds if and only if these worlds belong to different equivalence classes. If all pieces of evidence in the set $\{(u, a) \mid a \in \alpha(u)\}$ are eliminated, then the intruder is no longer able to distinguish epistemic worlds from these two equivalence classes. In what follows, we assume that only agent a can eliminate evidence (u, a) at a finite cost. The cost of eliminating all pieces of evidences in set $\{(u, a) \mid a \in \alpha(u)\}$ must be $\delta(u)$, but how exactly this cost is divided between all pieces of evidence (u, a) for various $a \in \alpha(u)$ is not important. For the sake of simplicity, we assume that the total cost $\delta(u)$ of removing pieces of evidence in set $\{(u, a) \mid a \in \alpha(u)\}$ is evenly divided between corresponding $a \in \alpha(u)$.

Definition 12. *For each $(u, a) \in \mathcal{E}$ and each $b \in \mathcal{A}$,*

$$\|(u,a)\|_b = \begin{cases} \dfrac{\delta(u)}{|\alpha(u)|}, & \textit{if } a = b, \\ +\infty, & \textit{otherwise,} \end{cases}$$

where $|\alpha(u)|$ denotes the cardinality of set $\alpha(u)$.

Cost function $\|(u, a)\|_b$ is well-defined because, by Definition 7, set $\alpha(u)$ is nonempty, and thus, $|\alpha(u)| \neq 0$.

Lemma 11. *If $\|(u, a)\|_B < +\infty$, then $a \in B$.*

Proof. The lemma follows from Definition 3 and Definition 12. ⊠

To conclude the definition of the canonical model, we next specify the semantics of the atomic propositions.

Definition 13. *For any atomic proposition p, let*

$$\pi(p) = \{w \in \mathcal{W} \mid p \in \sigma(w)\}.$$

The following three lemmas are variations of the lemmas commonly found in the proofs of completeness for various modal logics.

Lemma 12. *For any epistemic world $w \in \mathcal{W}$, if $\Box^c_A \varphi \notin \sigma(w)$, then there is $E \subseteq \mathcal{E}$ such that $\|E\|_A = c$ and there is a world $u \in \mathcal{W}$ such that $w \sim_{\mathcal{E} \setminus E} u$ and $\varphi \notin \sigma(u)$.*

Proof. We first show that the following set is consistent

$$X = \{\neg\varphi\} \cup \{\Box^d_B \psi \mid \Box^{d+c}_B \psi \in \sigma(w), A \subseteq B\} \cup \{\neg\Box^{d+c}_B \chi \mid \neg\Box^d_B \chi \in \sigma(w), A \subseteq B\}.$$

Assume the opposite. Thus, there must exist

$$\Box^{d_1+c}_{B_1}\psi_1, \dots, \Box^{d_n+c}_{B_n}\psi_n, \neg\Box^{d'_1}_{B'_1}\chi_1, \dots, \neg\Box^{d'_m}_{B'_m}\chi_m \in \sigma(w) \tag{3}$$

such that

$$\Box^{d_1}_{B_1}\psi_1, \dots, \Box^{d_n}_{B_n}\psi_n, \neg\Box^{d'_1+c}_{B'_1}\chi_1, \dots, \neg\Box^{d'_m+c}_{B'_m}\chi_m \vdash \varphi.$$

Hence, by the Deduction theorem for propositional logic,

$$\vdash \Box^{d_1}_{B_1}\psi_1 \to (\dots(\Box^{d_n}_{B_n}\psi_n \to (\neg\Box^{d'_1+c}_{B'_1}\chi_1 \to (\dots(\neg\Box^{d'_m+c}_{B'_m}\chi_m \to \varphi)\dots)))\dots).$$

Thus, by the Necessitation rule,

$$\vdash \Box^c_A(\Box^{d_1}_{B_1}\psi_1 \to (\dots(\Box^{d_n}_{B_n}\psi_n \to (\neg\Box^{d'_1+c}_{B'_1}\chi_1 \to (\dots(\neg\Box^{d'_m+c}_{B'_m}\chi_m \to \varphi)\dots)))\dots)).$$

By the Distributivity axiom and the Modus Ponens rule,

$$\Box^c_A\Box^{d_1}_{B_1}\psi_1 \quad \vdash \quad \Box^c_A(\Box^{d_2}_{B_2}\psi_2 \to (\dots \to (\neg\Box^{d'_m+c}_{B'_m}\chi_m \to \varphi)\dots)).$$

By repeating the previous step $(n+m-1)$ times,

$$\Box^c_A\Box^{d_1}_{B_1}\psi_1, \dots, \Box^c_A\Box^{d_n}_{B_n}\psi_n, \Box^c_A\neg\Box^{d'_1+c}_{B'_1}\chi_1, \dots, \Box^c_A\neg\Box^{d'_m+c}_{B'_m}\chi_m \vdash \Box^c_A\varphi.$$

By the Positive Introspection axiom applied n times,

$$\Box^{c+d_1}_{B_1}\psi_1, \dots, \Box^{c+d_n}_{B_n}\psi_n, \Box^c_A\neg\Box^{d'_1+c}_{B'_1}\chi_1, \dots, \Box^c_A\neg\Box^{d'_m+c}_{B'_m}\chi_m \vdash \Box^c_A\varphi.$$

By the Negative Introspection axiom applied m times,

$$\Box^{c+d_1}_{B_1}\psi_1, \dots, \Box^{c+d_n}_{B_n}\psi_n, \neg\Box^{d'_1}_{B'_1}\chi_1, \dots, \neg\Box^{d'_m}_{B'_m}\chi_m \vdash \Box^c_A\varphi.$$

Hence, $\sigma(w) \vdash \Box^c_A\varphi$, due to assumption (3). Thus, due to the maximality of $\sigma(w)$, we have $\Box^c_A\varphi \in \sigma(w)$, which contradicts the assumption of the lemma. Therefore, set X is consistent.

Let $\widehat{X}$ be a maximal consistent extension of X. Define u to be the sequence obtained by concatenating sequence $\langle (A,c), \widehat{X}\rangle$ to the end of sequence w. By Definition 7, we have $u \in W$.

Let $E = \{(u,a) \mid a \in \alpha(u)\}$. By Definition 4, Definition 3, Definition 12, and Definition 8,

$$\begin{aligned}\|E\|_A &= \|\{(u,a) \mid a \in A\}\|_A = \sum_{a\in A} \|(u,a)\|_A \\ &= \sum_{a\in A} \min_{b\in A} \|(u,a)\|_b = \sum_{a\in A} \frac{\delta(u)}{|\alpha(u)|} = \sum_{a\in A} \frac{c}{|A|} = c.\end{aligned}$$

Next we prove that $w \sim_{\mathcal{E}\setminus E} u$. It suffices to show that $w \sim_e u$ for any $e \in \mathcal{E}$ such that $e \notin E$. Consider any $e = (v,b) \in \mathcal{E} \setminus E$. Note that $v \neq u$ due to the choice of set E. By Lemma 8, statements $v \preceq w$ and $v \preceq u$ are equivalent. Hence, $w \sim_{(v,b)} u$ by Definition 10.

Finally, note that $\neg\varphi \in X$ due to the choice of set X. Thus, $\neg\varphi \in \widehat{X}$. Hence, $\varphi \notin \widehat{X}$ due to the consistency of set $\widehat{X}$. Therefore, $\varphi \notin \sigma(u)$ because $\sigma(u) = \widehat{X}$ by the choice of u. ⊠

Lemma 13. *For any two epistemic worlds $w_1, w_2 \in \mathcal{W}$, if $\Box^c_B\varphi \in \sigma(w_1)$, $\|E\|_B \leq c$, and $w_1 \sim_{\mathcal{E}\setminus E} w_2$, then $\varphi \in \sigma(w_2)$.*

Proof. Let $u = gcp(w_1, w_2)$. We start by proving three auxiliary claims.

Claim 1. *For every $(v,a) \in \mathcal{E}$, if either $u \prec v \preceq w_1$ or $u \prec v \preceq w_2$, then $(v,a) \in E$.*

Proof of Claim 1. Without loss of generality, it suffices to prove that $(v,a) \in E$, for every $(v,a) \in \mathcal{E}$ such that $u \prec v \preceq w_1$, see Figure 4. Indeed, since $u \prec v \preceq w_1$ and $u = gcp(w_1, w_2)$, we have $v \not\preceq w_2$. Thus, $w_1 \nsim_{(v,a)} w_2$ by Lemma 10. Therefore, $(v,a) \in E$ due to assumption $w_1 \sim_{\mathcal{E}\setminus E} w_2$ of the lemma. □

Claim 2. *For every $v \in \mathcal{W}$ and every $a \in \alpha(v)$, if either $u \prec v \preceq w_1$ or $u \prec v \preceq w_2$, then $a \in B$.*

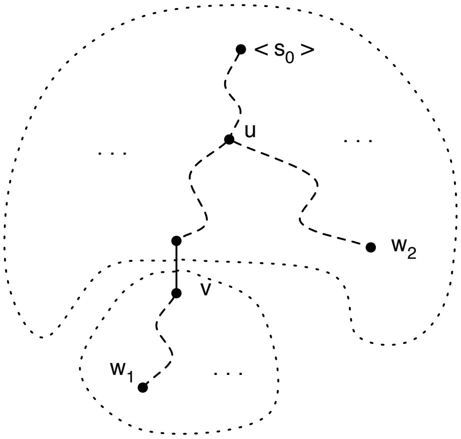

Figure 4: Towards the proof of Claim 1.

Proof of Claim 2. By Claim 1, we have $(v, a) \in E$. Thus, $\|(v, a)\|_B < +\infty$ due to assumption $\|E\|_B \leq c$ of the lemma. Therefore, $a \in B$ by Lemma 11. □

Claim 3. $\sum_{u \prec v \preceq w_1} \delta(v) + \sum_{u \prec v \preceq w_2} \delta(v) \leq c.$

Proof of Claim 3. By Definition 7, set $\alpha(v)$ is not empty for every $v \in \mathcal{W}$. Thus,

$$\sum_{u \prec v \preceq w_1} \delta(v) + \sum_{u \prec v \preceq w_2} \delta(v) = \sum_{u \prec v \preceq w_1} \sum_{a \in \alpha(v)} \frac{\delta(v)}{|\alpha(v)|} + \sum_{u \prec v \preceq w_2} \sum_{a \in \alpha(v)} \frac{\delta(v)}{|\alpha(v)|}.$$

At the same time, by Definition 12 and Claim 2,

$$\begin{aligned} &\sum_{u \prec v \preceq w_1} \sum_{a \in \alpha(v)} \frac{\delta(v)}{|\alpha(v)|} + \sum_{u \prec v \preceq w_2} \sum_{a \in \alpha(v)} \frac{\delta(v)}{|\alpha(v)|} \\ &= \sum_{u \prec v \preceq w_1} \sum_{a \in \alpha(v)} \|(v, a)\|_B + \sum_{u \prec v \preceq w_2} \sum_{a \in \alpha(v)} \|(v, a)\|_B. \end{aligned}$$

By Claim 1,

$$\sum_{u \prec v \preceq w_1} \sum_{a \in \alpha(v)} \|(v, a)\|_B + \sum_{u \prec v \preceq w_2} \sum_{a \in \alpha(v)} \|(v, a)\|_B \leq \sum_{e \in E} \|e\|_B.$$

Finally, by Definition 4 and assumption $\|E\|_B \leq c$ of the lemma,

$$\sum_{e\in E} \|e\|_B = \|E\|_B \leq c.$$

Therefore,

$$\sum_{u\prec v\preceq w_1} \delta(v) + \sum_{u\prec v\preceq w_2} \delta(v) \leq c.$$

$\square$

To finish the proof of the lemma, note that by Claim 2, Claim 3, and Corollary 2, assumption $\Box^c_B\varphi \in \sigma(w_1)$ implies

$$\Box_B^{c-\sum_{u\prec v\preceq w_1}\delta(v)}\varphi \in \sigma(u).$$

Hence, again by Claim 2, Claim 3, and Corollary 2,

$$\Box_B^{c-\sum_{u\prec v\preceq w_1}\delta(v)-\sum_{u\prec v\preceq w_2}\delta(v)}\varphi \in \sigma(w_2).$$

Thus, $\varphi \in \sigma(w_2)$ by the Truth axiom and the maximality of set $\sigma(w_2)$. This concludes the proof of Lemma 13. $\boxtimes$

Lemma 14. *$\varphi \in \sigma(w)$ if and only if $w \Vdash \varphi$, for each $w \in \mathcal{W}$ and each $\varphi \in \Phi$.*

Proof. We prove the statement of the lemma by induction on the structural complexity of formula φ.

Assume that formula φ is an atomic proposition p. Then, by Definition 13, statement $p \in \sigma(w)$ is equivalent to statement $w \in \pi(p)$. By Definition 6, statement $w \in \pi(p)$ is equivalent to $w \Vdash p$.

Cases when formula φ is an implication or a negation follow in the standard way from the induction hypothesis and the assumption of maximality and consistency of set $\sigma(w)$.

Suppose that φ is formula $\Box^c_A\psi$. ($\Rightarrow$) Let $\Box^c_A\psi \in \sigma(w)$. To prove that $w \Vdash \Box^c_A\psi$, consider any $E \subseteq \mathcal{E}$ such that $\|E\|_A \leq c$ and any $u \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E} u$. By Definition 6, it suffices to show that $u \Vdash \psi$, which is true by Lemma 13 and the induction hypothesis. ($\Leftarrow$) Assume that $\Box^c_A\psi \notin \sigma(w)$. Thus, by Lemma 12, there is subset $E \subseteq \mathcal{E}$ such that $\|E\|_A = c$ and there is world $u \in \mathcal{W}$ such that $w \sim_{\mathcal{E}\setminus E} u$ and $\psi \notin \sigma(u)$. Hence, by the induction hypothesis, $u \nVdash \psi$. Therefore, $w \nVdash \Box^c_A \psi$ by Definition 6. $\boxtimes$

We are now ready to state and to prove the completeness theorem for our logical system.

Theorem 1. *For any formula $\varphi \in \Phi$, if $w \Vdash \varphi$ for each epistemic world w of each Kripke model, then $\vdash \varphi$.*

Proof. Suppose that $\nvdash \varphi$. Let s_0 be any maximal consistent subset of Φ containing formula $\neg\varphi$. Consider canonical model $\mathcal{K}(s_0)$. By Lemma 14, $\langle s_0 \rangle \nVdash \varphi$. $\boxtimes$

6. Conclusion

In this article we proposed a logical framework for reasoning about the ability of a group of agents to protect privacy of their members at a given cost. The privacy protection is achieved by eliminating evidence and therefore hiding knowledge from a privacy intruder. Although throughout the article we have been assuming that the cost is monetary, the same system could be used to reason about many other types of cost such as billable time, supply resources, human resources, etc. The main technical results of this article are the soundness and the completeness theorems. Perhaps an even more important contribution of this work, however, is the proposed incorporation of the cost factor into privacy protection reasoning within the standard modal logic framework.

A natural extension of this work would be to consider protecting privacy from multiple privacy intruders. Note that most of the privacy in the real world is based on the assumption that the agents that have access to private information do not usually share this information between themselves. For example, patients usually assume that their doctors do not freely talk with patients' lawyers, bankers, and discount store managers. On rare occasion, however, such conversations happen and this often results in privacy violations. Thus, one can consider an extension of our logical system in which different intruders form coalitions similar to those in the setting of Pauly's Logic for Coalitional Power [10].

Finally, one can also develop a logical system for reasoning about the cost of *publicizing* the information rather than hiding it.

References

[1] H. Asher. *Polling and the Public: What Every Citizen Should Know: What Every Citizen Should Know*. SAGE Publications, 2011.

[2] Charles Duhigg. How companies learn your secrets. *New York Times*, February 16, 2012. www.nytimes.com/2012/02/19/magazine/shopping-habits.html.

[3] Ulle Endriss, Sarit Kraus, Jérôme Lang, and Michael Wooldridge. Designing incentives for boolean games. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 79–86. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[4] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about knowledge*. MIT Press, Cambridge, MA, 1995.

[5] Linda Holmes. When your data is currency, what does your privacy cost? *National Public Radio*, June 9, 2013. http://www.npr.org/blogs/monkeysee/2013/06/09/189857722/when-your-data-is-your-currency-what-does-your-privacy-cost.

[6] David Lazarus. Privacy price gouging, courtesy of phone companies. *Los Angeles Times*, April 14, 2014. www.latimes.com/business/la-fi-lazarus-20140415-column.html.

[7] Pavel Naumov and Jia Tao. Budget-constrained knowledge in multiagent systems. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 219–226. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[8] Pavel Naumov and Jia Tao. Logic of confidence. *Synthese*, 192(6):1821–1838, 2015.

[9] The Court of Justice of the European Union. Judgment of the Court (Grand Chamber) of 13 May 2014. Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González., 2014. Case: C-131/12.

[10] M. Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.

[11] Riva Richmond. How to fix (or kill) web data about you. *New York Times*, April 14, 2011. www.nytimes.com/2011/04/14/technology/personaltech/14basics.html.

[12] Sasha Romanosky and Alessandro Acquisti. Privacy costs and personal data protection: Economic and legal perspectives. *Berkeley Technology Law Journal*, 24:1061, 2009.

[13] Henrik Sahlqvist. Completeness and correspondence in the first and second order semantics for modal logic. *Studies in Logic and the Foundations of Mathematics*, 82:110–143, 1975. (Proc. of the 3rd Scandinavial Logic Symposium, Uppsala, 1973).

[14] Adam Shostack. Paying for privacy: Consumers and infrastructures. In *2nd Annual Workshop on Economics and Information Security-WEIS*, volume 3, 2003.

[15] Johan van Benthem and Eric Pacuit. Dynamic logics of evidence-based beliefs. *Studia Logica*, 99(1-3):61–92, 2011.

[16] Hans van Ditmarsch, Andreas Herzig, Jérôme Lang, and Pierre Marquis. Introspective forgetting. *Synthese*, 169(2):405–423, 2009.

[17] Hal R. Varian. Economic aspects of personal privacy. In William H. Lehr and Lorenzo Maria Pupillo, editors, *Internet Policy and Economics*, pages 101–109. Springer US, 2009.

[18] Armgard von Reden. The costs and benefits of privacy. In *The 26th International Conference on Privacy and Personal Data Protection*, 2004.

[19] Yan Zhang and Yi Zhou. Knowledge forgetting: Properties and applications. *Artificial Intelligence*, 173(16):1525–1537, 2009.