

## Logic of Confidence

Pavel Naumov · Jia Tao

the date of receipt and acceptance should be inserted later

**Abstract** The article studies knowledge in multiagent systems where data available to the agents may have small errors. To reason about such uncertain knowledge, a formal semantics is introduced in which indistinguishability relations, commonly used in the semantics for epistemic logic S5, are replaced with metrics to capture how much two epistemic worlds are different from an agent’s point of view. The main result is a logical system sound and complete with respect to the proposed semantics.

**Keywords** Uncertainty · Confidence of Knowledge · Axiomatic System

### 1 Introduction

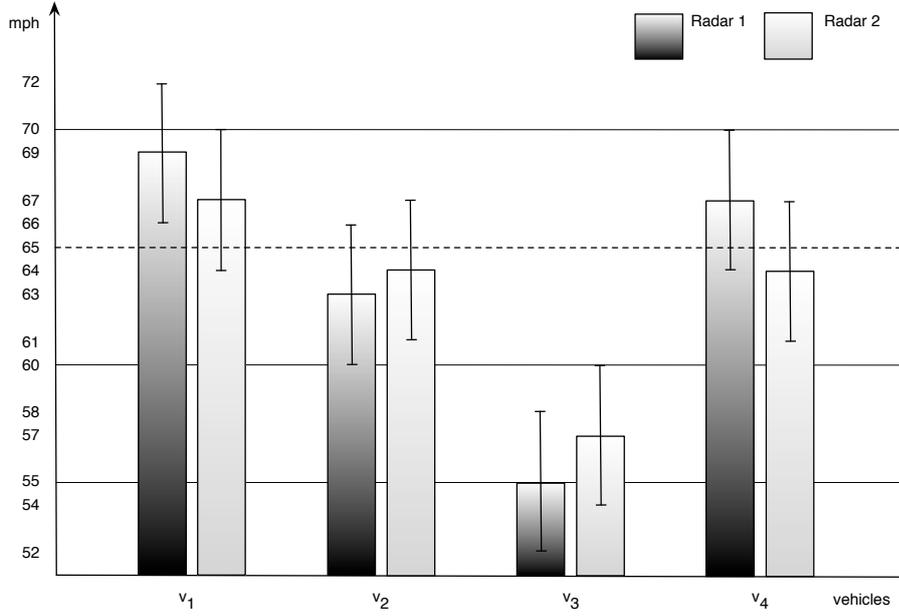
Uncertainty about information in multiagent systems often comes from the imprecision in measurements or small errors in data. Consider an example of the radar data about four different vehicles depicted in Figure 1. The data is reported by two different police officers. The first police officer,  $p_1$ , observes vehicle  $v_1$  going with speed 69 miles per hour (mph). Assuming that the precision of the radar is  $\pm 3$  mph, officer  $p_1$  concludes that the actual speed of vehicle  $v_1$ , in miles per hour, is in the interval  $[66, 72]$ . Such an interval is often referred to as a *confidence interval*. Since driving a vehicle at any speed in this interval exceeds 65 mph speed limit, officer  $p_1$  concludes that the driver of vehicle  $v_1$  breaks the traffic law, as long as the precision of the radar is indeed  $\pm 3$  mph. We write it formally as

$$\Box_{p_1}^3 (\text{“Vehicle } v_1 \text{ is speeding.”}).$$

---

Pavel Naumov  
McDaniel College, Westminster, Maryland 21157, USA  
E-mail: pavel@pavelnaumov.com

Jia Tao  
Bryn Mawr College, Bryn Mawr, Pennsylvania 19010, USA  
E-mail: jtiao@cs.brynmawr.edu



**Fig. 1** Confidence intervals for data obtained by officers  $p_1$  (Radar 1) and  $p_2$  (Radar 2).

At the same time, the second police officer,  $p_2$ , observes the radar reading of 67 mph and, again given the radar precision of  $\pm 3$  mph, can not be completely sure that the traffic law is violated:

$$\neg \square_{p_2}^3 (\text{“Vehicle } v_1 \text{ is speeding.”}).$$

However, if officer  $p_2$  has some reason to believe that the precision of the measurement is actually  $\pm 1$  mph, then she can make the same conclusion as officer  $p_1$ :

$$\square_{p_2}^1 (\text{“Vehicle } v_1 \text{ is speeding.”}).$$

Neither of officers  $p_1$  and  $p_2$  knows for sure if vehicles  $v_2$  and  $v_4$  are speeding, but both of them are confident that vehicle  $v_3$  is not speeding. Furthermore, since officer  $p_1$  knows that the actual speed of vehicle  $v_3$  is in the interval  $[52, 58]$ , she knows that the reading of officer  $p_2$ 's radar is in the interval  $[49, 61]$ . Thus, she infers that the confidence interval of officer  $p_2$  is a subinterval of  $[46, 64]$ . Therefore, she knows that officer  $p_2$  knows that vehicle  $v_3$  is not speeding:

$$\square_{p_1}^3 \square_{p_2}^3 \neg (\text{“Vehicle } v_3 \text{ is speeding.”}).$$

At the same time, officer  $p_2$ , reasoning in a similar way, can predict that the confidence interval of officer  $p_1$  is a subinterval of  $[48, 66]$ . Thus, she is not able to conclude that  $p_1$  knows that vehicle  $v_3$  is speeding:

$$\neg \square_{p_2}^3 \square_{p_1}^3 \neg (\text{“Vehicle } v_3 \text{ is speeding.”}).$$

As another example, imagine a spy  $s$  getting a secret message from an intelligence agency through a shortwave number station. The spy is expecting to get a binary code representing one of three possible instructions listed in Figure 2. Note that any two binary codes in this table differ in at least four digits. As a result, if only one digit is corrupted due to the radio noise, the agent will still be able to recover the instruction. For example, if the agent gets code 1010011, then it is understood that the first instruction in Figure 2 is sent rather than any of the other two.

Binary Code	Instruction
1010010	“meet the carrier at the train station”
0111000	“meet the carrier in the park”
1101101	“escape, you have been exposed”

**Fig. 2** Spy Instruction Codes.

Assume that due to extra radio noise the spy  $s$  misinterprets two digits in the transmission and gets code 1110000. If the spy  $s$  correctly assumes that at most two digits are corrupted due to the noise, he concludes that one of the first two instructions in Figure 2 is sent, but he is not able to find out which one of them is sent:

$$\begin{aligned} & \Box_s^2(\text{“the carrier is arriving”}), \\ & \neg\Box_s^2(\text{“the carrier will be at the train station”}), \\ & \neg\Box_s^2(\text{“the carrier will be in the park”}). \end{aligned}$$

Suppose that the same transmission is intercepted by a counterintelligence agency  $a$  that not only is in possession of the instruction codes shown in Figure 2, but also has a better equipment that is expected to misinterpret at most one digit even if the radio transmission is extra noisy. If the agency receives code 1111000, then it knows that the carrier will be in the park:

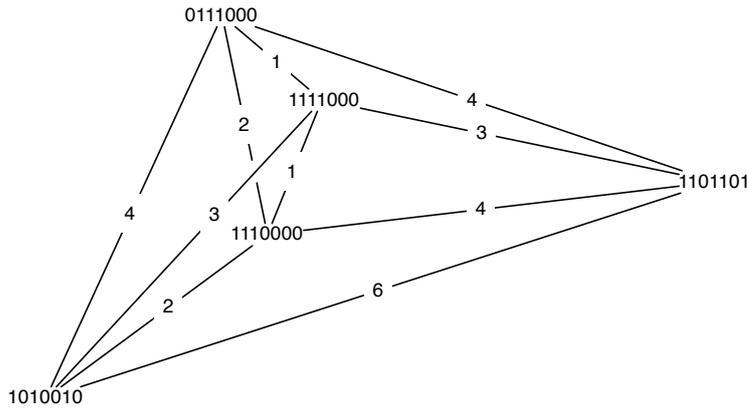
$$\Box_a^1(\text{“the carrier will be in the park”}).$$

Furthermore, if the spy knows about technical capabilities of the agency, then he should be able to conclude that the agency knows not only about the carrier’s arrival, but also about the planned meeting place:

$$\begin{aligned} & \Box_s^2(\Box_a^1(\text{“the carrier will be in the park”})) \\ & \vee \Box_s^1(\Box_a^1(\text{“the carrier will be at the station”})). \end{aligned}$$

An easy way (see Figure 3) to visualize this situation is to draw a picture showing each binary string as a point. For any two strings, the “distance” between them is the number of positions at which digits of these two strings differ. This distance is known as Hamming distance [5].

In our first example, the data about a vehicle shown in Figure 1, together with its actual speed, could be thought of as an “epistemic world” in the traditional epistemic logic sense. In the second example, the epistemic world



**Fig. 3** Hamming Distances between Binary Strings.

is the combination of observations by the spy and by the counterintelligence agency and the actual message transmitted by the intelligence agency. In either of these two cases, agents have an agent-specific way to measure the distance between different epistemic worlds. In the first example, the distance between epistemic words is the absolute value of the difference between observations by the appropriate agent, which is also called Euclidean distance. In the second example, the distance is the Hamming distance between such observations.

In mathematics, a set  $W$  with a distance function  $d$  on  $W \times W$  that satisfies the following three properties

1. Identity of Indiscernibles:  $d(u, v) = 0$  if and only if  $u = v$
2. Symmetry:  $d(u, v) = d(v, u)$
3. Triangle Inequality:  $d(u, v) \leq d(u, w) + d(w, v)$

is called a “metric space”. In this article we develop a logical system for reasoning about confidence modality  $\Box_a^c$  for an arbitrary metric space. We call this modality “confidence” after the commonly used term “confidence interval” in situations when the metric space is the set of real numbers  $\mathbb{R}$  with Euclidean distance on it, as in our first example. Superscript  $c$  is the half of the length of the confidence interval or, in the case of an arbitrary metric space, the radius of the “confidence ball”. The axioms of the system closely resemble axioms of multiagent epistemic logic S5 with confidence superscript  $c$  added (see Section 3). Our main result is the completeness theorem for this logical system with respect to a formally defined semantics.

Our use of confidence intervals to capture uncertainty is closely related to Halpern and Shoham’s work on interval temporal logic [4]. The modality in their system, however, is temporal and does not have the same epistemic meaning as ours. Due to their focus on temporal reasoning, they do not consider the setting of an arbitrary metric space. They analyse the complexity of the interval logic, but give no explicit axiomatization.

Milnikel [8] proposed a version of justification logic [1] with uncertain justifications. His system studies the *degrees* up to which an agent is confident that certain evidence justifies the formula. The difference between our and his approaches manifests itself most explicitly in the modal distributivity axiom. In our system (see Section 3) the distributivity axiom is:  $\Box_a^c(\varphi \rightarrow \psi) \rightarrow (\Box_a^c\varphi \rightarrow \Box_a^c\psi)$ . Informally, it states that if agent  $a$  is confident in  $\varphi \rightarrow \psi$ , assuming precision of measurements  $\pm c$ , and she is also confident in  $\varphi$ , assuming the same precision of measurements, then she is confident in  $\psi$ , again with the same precision. The corresponding axiom in Milnikel’s system [8] could be written as  $\Box_a^c(\varphi \rightarrow \psi) \rightarrow (\Box_b^d\varphi \rightarrow \Box_{a,b}^{c \cdot d}\psi)$ , where  $c \cdot d$  is the product of two real numbers representing agent’s degrees of confidence that the combination of justifications  $a$  and  $b$  supports claim  $\psi$ .

Our work at its core is very much connected to Moss and Parikh’s work on topological reasoning and epistemic logic [9]. Similar to ours, their work started with confidence intervals, although they did not use this term explicitly. The same example about the accuracy of a police officer’s knowledge of the speed of a car was used to introduce the topic. They chose, however, a much more general setting than ours. Instead of generalizing from confidence intervals on a number line to a “ball” in an arbitrary metric space as we do, they went one step further by interpreting it as an open set in a topological space. The smaller such a set is the more precise the knowledge is. As a result, their logical system can reason about different levels of precision without assigning a numerical value to this precision. They also separated the knowledge modality into two different modal operators: one representing the knowledge and the other representing the precision of the knowledge (or, to use their terminology, “the effort”).

Finally, our work could also be thought of as a study of spatial reasoning in an arbitrary metric space. The satisfiability relation  $w \Vdash \Box_a^c\varphi$ , from such a perspective, could be interpreted as “statement  $\varphi$  is true in  $c$ -neighbourhood of point  $w$  under the metric associated with agent  $a$ ”. This relates our system to other modal logics for spatial reasoning [12, 3].

The rest of the article is structured as follows. Section 2 formally specifies syntax and semantics of the logical system. Section 3 introduces its axioms and inference rules. The soundness of the system with respect to the formal semantics is shown in Section 4. The proof of the completeness is divided into two parts. Section 5 considers metric spaces in which distances between points could be infinite and proves the completeness theorem with respect to the class of confidence models based on such metric spaces. The proof employs the “unravelling” [11] technique from modal logic model theory. Section 6 uses truncation of metric spaces to prove the completeness for the class of confidence models based on the metric spaces with finite distances. Section 7 discusses possible generalizations of our logic to group confidence as well as its connections to standard Kripke models and judgement aggregation.

## 2 Syntax and Semantics

In this section we formally define syntax and semantics of our logical system. We assume a fixed set  $\mathcal{P}$  of atomic propositions and a fixed set  $\mathcal{A}$  of agents.

**Definition 1** Let  $\Phi$  be the smallest set of formulas such that

1.  $\mathcal{P} \subseteq \Phi$ ,
2. if  $\varphi, \psi \in \Phi$ , then  $\neg\varphi \in \Phi$  and  $\varphi \rightarrow \psi \in \Phi$ ,
3. if  $\varphi \in \Phi$ , then  $\Box_a^c \varphi \in \Phi$ , for each  $a \in \mathcal{A}$ , and each real number  $c \geq 0$ .

Below is the standard definition of a metric space [10]. Hamming distance [5], Manhattan distance [7], shortest path distance, and Euclidean space distance are standard examples of such metric spaces.

**Definition 2** A finite metric space is a pair  $(W, d)$  such that  $W$  is a set and  $d$  is a distance function between two elements of  $W$  whose value is a non-negative real number. Additionally, function  $d$  must satisfy the following properties for all  $u, v, w \in W$ :

1. Identity of Indiscernibles:  $d(u, v) = 0$  if and only if  $u = v$ ,
2. Symmetry:  $d(u, v) = d(v, u)$ ,
3. Triangle Inequality:  $d(u, v) \leq d(u, w) + d(w, v)$ .

In the above definition, as in some of the literature on metric spaces [2], we use the term *finite* metric space. In Section 5, we consider a more general definition of a metric space in which distances can be infinite.

**Definition 3** A finite confidence model is a tuple  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$ , such that  $(W, d_a)$  is a finite metric space for each  $a \in \mathcal{A}$  and  $\pi$  is a function from atomic propositions into subsets of  $W$ . Elements of set  $W$  are called “epistemic worlds”.

In Section 7, we discuss relation between finite confidence models and standard Kripke models used in epistemic logic.

**Definition 4** For any world  $w$  of a finite confidence model  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$  and any formula  $\varphi \in \Phi$ , we define satisfiability relation  $w \Vdash \varphi$  as follows:

1.  $w \Vdash p$  if  $w \in \pi(p)$ , where  $p \in \mathcal{P}$ ,
2.  $w \Vdash \neg\psi$  if  $w \not\Vdash \psi$ , where  $\psi \in \Phi$ ,
3.  $w \Vdash \psi_1 \rightarrow \psi_2$  if  $w \not\Vdash \psi_1$  or  $w \Vdash \psi_2$ , where  $\psi_1, \psi_2 \in \Phi$ ,
4.  $w \Vdash \Box_a^c \psi$  if  $u \Vdash \psi$  for each epistemic world  $u \in W$  such that  $d_a(w, u) \leq c$ , where  $\psi \in \Phi$ ,  $a \in \mathcal{A}$ , and  $c \geq 0$ .

## 3 Axioms and Inference Rules

Our logical system, in addition to propositional tautologies in language  $\Phi$ , contains the following axioms:

1. Zero Confidence:  $\varphi \rightarrow \Box_a^0 \varphi$ ,
2. Truth:  $\Box_a^c \varphi \rightarrow \varphi$ ,
3. Positive Introspection:  $\Box_a^{c+d} \varphi \rightarrow \Box_a^c \Box_a^d \varphi$ ,
4. Negative Introspection:  $\neg \Box_a^c \varphi \rightarrow \Box_a^d \neg \Box_a^{c+d} \varphi$ ,
5. Distributivity:  $\Box_a^c (\varphi \rightarrow \psi) \rightarrow (\Box_a^c \varphi \rightarrow \Box_a^c \psi)$ .

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

$$\frac{\varphi, \varphi \rightarrow \psi}{\psi} \qquad \frac{\varphi}{\Box_a^c \varphi}$$

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

#### 4 Soundness

In this section we establish the soundness of our logical system with respect to the semantics given in Definition 2. The soundness of propositional tautologies and of Modus Ponens inference rule is straightforward. Below we prove the soundness of each remaining axiom and of the Necessitation inference rule as a separate lemma. We assume that  $w$  is an arbitrary epistemic world of a finite confidence model  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$ .

**Lemma 1 (Zero Confidence)** *For any formula  $\varphi \in \Phi$  and any  $a \in \mathcal{A}$ , if  $w \Vdash \varphi$ , then  $w \Vdash \Box_a^0 \varphi$ .*

*Proof* We need to prove that  $u \Vdash \varphi$  for each  $u \in W$  such that  $d_a(w, u) = 0$ . By Identity of Indiscernibles property of metric spaces (see Definition 2), equality  $d_a(w, u) = 0$  implies  $w = u$ . Hence,  $u \Vdash \varphi$  due to the assumption  $w \Vdash \varphi$ .  $\square$

**Lemma 2 (Truth)** *For any formula  $\varphi \in \Phi$ , any  $a \in \mathcal{A}$ , and any real number  $c \geq 0$ , if  $w \Vdash \Box_a^c \varphi$ , then  $w \Vdash \varphi$ .*

*Proof* Suppose that  $w \Vdash \Box_a^c \varphi$ . By Identity of Indiscernibles property of metric spaces (see Definition 2),  $d_a(w, w) = 0$ . Thus,  $d_a(w, w) \leq c$ . Hence, by Definition 4,  $w \Vdash \varphi$ .  $\square$

**Lemma 3 (Positive Introspection)** *For any formula  $\varphi \in \Phi$ , any  $a \in \mathcal{A}$ , and any real numbers  $c, d \geq 0$ , if  $w \Vdash \Box_a^{c+d} \varphi$ , then  $w \Vdash \Box_a^c \Box_a^d \varphi$ .*

*Proof* Consider any  $u \in W$  such that  $d_a(w, u) \leq c$ . By Definition 4, it suffices to show that  $u \Vdash \Box_a^d \varphi$ . Indeed, let  $v \in W$  be any epistemic world such that  $d_a(u, v) \leq d$ . We need to prove that  $v \Vdash \varphi$ . By Triangle Inequality,  $d_a(w, v) \leq d_a(w, u) + d_a(u, v) \leq c + d$ . Thus, due to assumption  $w \Vdash \Box_a^{c+d} \varphi$  and Definition 4,  $v \Vdash \varphi$ .  $\square$

**Lemma 4 (Negative Introspection)** *For any formula  $\varphi \in \Phi$ , any  $a \in \mathcal{A}$ , and any real numbers  $c, d \geq 0$ , if  $w \Vdash \neg \Box_a^c \varphi$ , then  $w \Vdash \Box_a^d \neg \Box_a^{c+d} \varphi$ .*

*Proof* By Definition 4, assumption  $w \Vdash \neg \Box_a^c \varphi$  implies the existence of an epistemic world  $u \in W$  such that  $d_a(w, u) \leq c$  and  $u \not\Vdash \varphi$ . To prove  $w \Vdash \Box_a^d \neg \Box_a^{c+d} \varphi$ , consider any  $v \in W$  such that  $d_a(w, v) \leq d$ . It suffices to show that  $v \not\Vdash \Box_a^{c+d} \varphi$ . Indeed, by Triangle Inequality,  $d_a(v, u) \leq d_a(v, w) + d_a(w, u) \leq d + c$ . Additionally,  $u \not\Vdash \varphi$  due to the choice of world  $u$ . Therefore, by Definition 4,  $v \not\Vdash \Box_a^{c+d} \varphi$ .  $\square$

**Lemma 5 (Distributivity)** *For any formulas  $\varphi, \psi \in \Phi$ , any  $a \in \mathcal{A}$ , and any real number  $c \geq 0$ , if  $w \Vdash \Box_a^c(\varphi \rightarrow \psi)$  and  $w \Vdash \Box_a^c \varphi$ , then  $w \Vdash \Box_a^c \psi$ .*

*Proof* Consider any  $u \in W$  such that  $d_a(w, u) \leq c$ . It suffices to show that  $u \Vdash \psi$ . Indeed,  $u \Vdash \varphi \rightarrow \psi$  and  $u \Vdash \varphi$  due to Definition 4 and assumptions  $w \Vdash \Box_a^c(\varphi \rightarrow \psi)$  and  $w \Vdash \Box_a^c \varphi$  respectively. Therefore, again by Definition 4,  $u \Vdash \psi$ .  $\square$

**Lemma 6 (Necessitation)** *For any formula  $\varphi \in \Phi$ , if  $u \Vdash \varphi$  for each world  $u$  of each finite confidence model, then  $u \Vdash \Box_a^c \varphi$  for each world  $u$  of each finite confidence model, where  $a \in \mathcal{A}$  and  $c$  is a non-negative real number.*

*Proof* Suppose that  $v \in W$  is such that  $d_a(u, v)$ . It suffices to show that  $v \Vdash \varphi$ , which is true due to the assumption of the lemma.  $\square$

## 5 Completeness for Confidence Models

The goal of this and the next section is to prove the completeness theorem for our logical system with respect to finite confidence models specified in Definition 3. This result is obtained in two steps. In this section, we generalize *finite* confidence models to confidence models and prove the completeness of our system with respect to this class of more general models. The proof employs the “unravelling” [11] technique from modal logic model theory. In the next section, we show how to convert a confidence model to a finite confidence model, using a technique from metric space theory sometimes (for example, in [6]) called “truncation”.

### 5.1 Metric Space and Confidence Models

In Definition 2, we require each distance to be a finite real number. It is sometimes more convenient to use a more general notion of a metric space with possibly infinite distances. Following [2], we refer to the more general notion as just a *metric space* as opposed to the earlier discussed notion of a *finite* metric space.

**Definition 5** A metric space is a pair  $(W, d)$  such that  $W$  is a set and  $d$  is a distance function between two elements of  $W$  whose value is a non-negative real number or  $\infty$ . Additionally, function  $d$  must satisfy the following properties for all  $u, v, w \in W$ :

1. Identity of Indiscernibles:  $d(u, v) = 0$  if and only if  $u = v$ ,
2. Symmetry:  $d(u, v) = d(v, u)$ ,
3. Triangle Inequality:  $d(u, v) \leq d(u, w) + d(w, v)$ .

In the above definition we assume that  $\infty + \infty = \infty$  and  $\infty + r = \infty$ , where  $r$  is any non-negative real number.

We define a more general notion of a confidence model using metric spaces in the same way finite confidence models are specified in Definition 3 using finite metric spaces.

**Definition 6** A confidence model is a tuple  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$ , such that  $(W, d_a)$  is a metric space for each  $a \in \mathcal{A}$  and  $\pi$  is a function from atomic propositions into subsets of  $W$ . Elements of set  $W$  are called “epistemic worlds”.

The satisfiability relation for confidence models is defined below in the same way as in the case of finite confidence models (Definition 4).

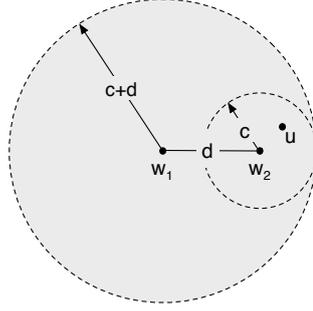
**Definition 7** For any world  $w$  of an confidence model  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$  and any formula  $\varphi \in \mathcal{F}$ , we define satisfiability relation  $w \Vdash \varphi$  as follows:

1.  $w \Vdash p$  if  $w \in \pi(p)$ , where  $p \in \mathcal{P}$ ,
2.  $w \Vdash \neg\psi$  if  $w \not\Vdash \psi$ , where  $\psi \in \mathcal{F}$ ,
3.  $w \Vdash \psi_1 \rightarrow \psi_2$  if  $w \not\Vdash \psi_1$  or  $w \Vdash \psi_2$ , where  $\psi_1, \psi_2 \in \mathcal{F}$ ,
4.  $w \Vdash \Box_a^c \psi$  if  $u \Vdash \psi$  for each epistemic world  $u \in W$  such that  $d_a(w, u) \leq c$ , where  $\psi \in \mathcal{F}$ ,  $a \in \mathcal{A}$ , and  $c \geq 0$ .

## 5.2 Canonical Confidence Model

In this section we modify the standard canonical epistemic model construction to produce a canonical confidence model. The epistemic worlds in the canonical epistemic model are usually defined as maximal consistent sets of formulas. Two such sets  $w_1$  and  $w_2$  are indistinguishable to an agent  $a$  when  $\Box_a \varphi \in w_1$  if and only if  $\Box_a \varphi \in w_2$  for each formula  $\varphi$ . To construct a canonical *confidence* model, instead of indistinguishability relation, we need to define the distance between maximal consistent sets  $w_1$  and  $w_2$  under the metric associated with any given agent  $a$ . We “reverse engineer” this definition by assuming that the distance between  $w_1$  and  $w_2$  is  $d$  (see Figure 4) and by studying the relation between formulas in sets  $w_1$  and  $w_2$ .

Suppose that  $\Box_a^{c+d} \varphi \in w_1$ . We would like this to imply that in our canonical model  $w_1 \Vdash \Box_a^{c+d} \varphi$ . Then, by Definition 7,  $u \Vdash \varphi$  for each world  $u$  inside the ball of radius  $c + d$  around epistemic world  $w_1$ . Hence, by the Triangle Inequality,  $u \Vdash \varphi$  for each  $u$  inside the ball of radius  $c$  around epistemic world  $w_2$  (see Figure 4). In other words,  $w_2 \Vdash \Box_a^c \varphi$ , which, in our canonical



**Fig. 4** Toward defining distance between epistemic worlds  $w_1$  and  $w_2$ .

model, we would like to imply  $\Box_a^c \varphi \in w_2$ . Thus, due to the symmetry of the distance between sets  $w_1$  and  $w_2$ , we have “reverse engineered” the following two properties that connect formulas in these two sets:

1. if  $\Box_a^{c+d} \varphi \in w_1$ , then  $\Box_a^c \varphi \in w_2$ ,
2. if  $\Box_a^{c+d} \varphi \in w_2$ , then  $\Box_a^c \varphi \in w_1$ .

Our first idea was to define distance  $d_a(w_1, w_2)$  as the minimal  $d$  that satisfies the two properties above. Unfortunately, if sets  $w_1$  and  $w_2$  are infinite, then such minimal  $d$  might not exist. As a result, we enhance the construction of the canonical confidence model  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$  by employing the “unraveling” technique [11]. In the enhanced construction, the two properties above correspond to conditions (a) and (b) in item 2. of Definition 8 below.

Throughout this section, we assume that  $s_0$  is any fixed maximal consistent subset of  $\Phi$ . The distance associated with some of the agents in the defined below model might not be finite.

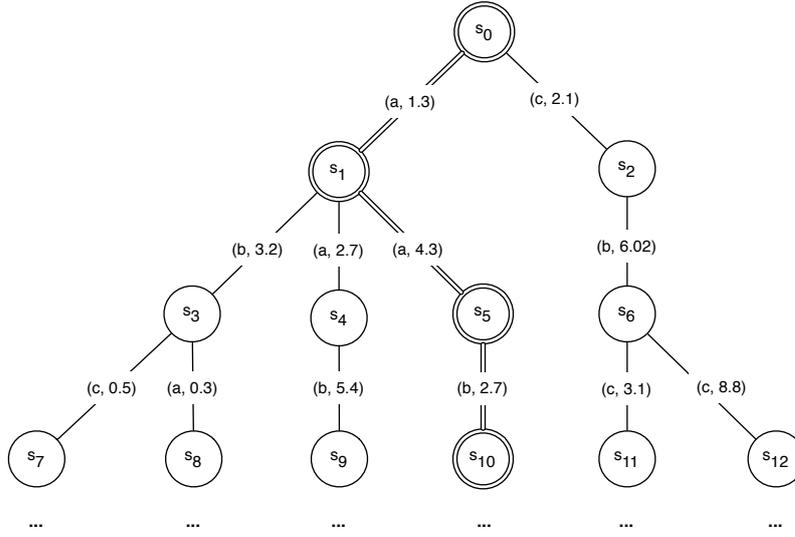
**Definition 8** Let the set of epistemic worlds  $W$  be the set of all sequences

$$\langle s_0, (a_1, d_1), s_1, (a_2, d_2), s_2, \dots, (a_n, d_n), s_n \rangle$$

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

1.  $s_i$  is a maximal consistent subset of  $\Phi$ ,
2. agent  $a_i \in \mathcal{A}$  and real number  $d_i > 0$  are such that the following two conditions are satisfied for each  $c \geq 0$  and each  $\varphi \in \Phi$ :
  - (a) if  $\Box_a^{c+d_i} \varphi \in s_{i-1}$ , then  $\Box_a^c \varphi \in s_i$ ,
  - (b) if  $\Box_a^{c+d_i} \varphi \in s_i$ , then  $\Box_a^c \varphi \in s_{i-1}$ .

Informally, epistemic worlds of the canonical model can be viewed as finite paths (see Figure 5) in the infinite tree whose vertices are maximal consistent sets and whose edges are labeled by pairs  $(a, d)$ , where  $a \in \mathcal{A}$  and  $d > 0$ . For any epistemic world  $w = \langle s_0, (a_1, d_1), s_1, (a_2, d_2), s_2, \dots, (a_n, d_n), s_n \rangle$ , by  $end(w)$  we mean maximal consistent set  $s_n$ .



**Fig. 5** Sequence  $\langle s_0, (a, 1.3), s_1, (a, 4.3), s_5, (b, 2.7), s_{10} \rangle$  is an example of an epistemic world of the canonical confidence model.

**Definition 9** For any two epistemic worlds  $w, u \in W$ , let  $w \preceq u$  denote that sequence  $w$  is a prefix of sequence  $u$ .

For example,  $\langle s_0, (a, 1.3), s_1 \rangle \preceq \langle s_0, (a, 1.3), s_1, (a, 4.3), s_5, (b, 2.7), s_{10} \rangle$  for the epistemic worlds depicted in Figure 5.

**Definition 10** The longest common prefix  $lcp(u, v)$  of two epistemic worlds  $u, v \in W$  is the longest sequence  $w \in W$  such that  $w \preceq u$  and  $w \preceq v$ .

For the example in Figure 5, the longest common prefix of epistemic worlds  $\langle s_0, (a, 1.3), s_1, (b, 3.2), s_3, (c, 0.5), s_7 \rangle$  and  $\langle s_0, (a, 1.3), s_1, (b, 3.2), s_3, (a, 0.3), s_8 \rangle$  is epistemic world  $\langle s_0, (a, 1.3), s_1, (b, 3.2), s_3 \rangle$ . Next we define the distance  $d_a(w, u)$  between two worlds  $w, u \in W$ . First, we consider the special case when  $w \preceq u$ .

**Definition 11** For any epistemic worlds  $w, u \in W$ , if

$$u = \langle w, (a_{k+1}, d_{k+1}), s_{k+1}, (a_{k+2}, d_{k+2}), s_{k+2}, \dots, (a_n, d_n), s_n \rangle,$$

then

$$d_a(w, u) = \begin{cases} \sum_{i=k+1}^n d_i, & \text{if } a_i = a \text{ for all } k+1 \leq i \leq n, \\ \infty, & \text{otherwise.} \end{cases}$$

For example, for the epistemic worlds depicted in Figure 5,

$$\begin{aligned} d_a(\langle s_0 \rangle, \langle s_0, (a, 1.3), s_1, (a, 4.3), s_5 \rangle) &= 1.3 + 4.3 = 5.6, \\ d_a(\langle s_0 \rangle, \langle s_0, (a, 1.3), s_1, (a, 4.3), s_5, (b, 2.7), s_{10} \rangle) &= \infty. \end{aligned}$$

We define distance between two arbitrary epistemic worlds as follows:

**Definition 12** For any two epistemic worlds  $u, v \in W$ , the distance between  $u$  and  $v$  is:  $d_a(u, v) = d_a(lcp(u, v), u) + d_a(lcp(u, v), v)$ .

For example, for the epistemic worlds depicted in Figure 5,

$$\begin{aligned} d_a(\langle s_0, (a, 1.3), s_1, (a, 2.7), s_4 \rangle, \langle s_0, (a, 1.3), s_1, (a, 4.3), s_5 \rangle) &= 2.7 + 4.3 = 7.0, \\ d_a(\langle s_0, (a, 1.3), s_1, (b, 3.2), s_3 \rangle, \langle s_0, (a, 1.3), s_1, (a, 4.3), s_5 \rangle) &= \infty. \end{aligned}$$

**Lemma 7** For any two worlds  $w, u \in W$  such that  $w \preceq u$  and  $d_a(w, u) \leq c$ ,

1. if  $\Box_a^c \varphi \in \text{end}(w)$ , then  $\Box_a^{c-d_a(w, u)} \varphi \in \text{end}(u)$ ,
2. if  $\Box_a^c \varphi \in \text{end}(u)$ , then  $\Box_a^{c-d_a(w, u)} \varphi \in \text{end}(w)$ .

*Proof* We prove by induction on the length of sequence  $u$ .

*Base Case:*  $w = u$ . Then the required follows from  $d_a(w, u) = 0$ .

*Induction Step:* The induction step for parts 1 and 2 of the lemma follows, respectively, from conditions a) and b) of part 2 of Definition 8.  $\square$

Shortest path distance between any two points of an arbitrary weighted graph satisfies all conditions in Definition 2 and, thus, specifies metric between vertices of the graph. By definition, a tree is a graph that has exactly one path without self-intersections between any two vertices. Thus, the length of such a path is a metric on vertices of an arbitrary weighted tree. One can see that function  $d_a(w, u)$  is equal to the length of the shortest path between vertices  $\text{end}(w)$  and  $\text{end}(v)$  on the tree partially depicted in Figure 5. This implies that  $d_a$  is a metric on the set of epistemic worlds  $W$ . However, since we have never formally defined the class of trees whose example is depicted in Figure 5, our proof below, instead of using shortest distance argument, directly refers to Definition 12.

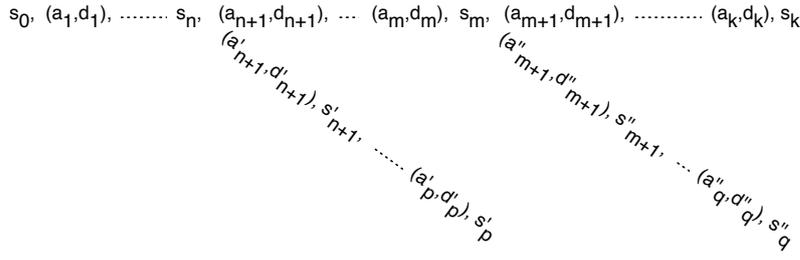
**Lemma 8** The pair  $(W, d_a)$  is a metric space for each  $a \in \mathcal{A}$ .

*Proof* Identity of Indiscernibles property follows from our assumption  $d_i > 0$  in Definition 8. Symmetry property follows from the observation  $lcp(u, v) = lcp(v, u)$  and Definition 12. To prove Triangle Inequality, we need to show that  $d_a(u, w) \leq d_a(u, v) + d_a(v, w)$  for each  $u, v, w \in W$ . Note that out of three sequences,  $u, v$ , and  $w$ , one might “split off” earlier than the other two (see Figure 6).

*Case I:* Assume that either  $w$  or  $u$  splits off first. Without loss of generality, suppose that  $w$  splits off first. Then (see Figure 6), let

$$\begin{aligned} u &= s_0, (a_1, d_1), \dots, s_n, (a_{n+1}, d_{n+1}), \dots, (a_m, d_m), s_m, (a_{m+1}, d_{m+1}), \dots, s_k, \\ w &= s_0, (a_1, d_1), \dots, s_n, (a'_{n+1}, d'_{n+1}), \dots, (a'_p, d'_p), s'_p, \\ v &= s_0, (a_1, d_1), \dots, s_n, (a_{n+1}, d_{n+1}), \dots, (a_m, d_m), s_m, (a''_{m+1}, d''_{m+1}), \dots, s''_q. \end{aligned}$$

If at least one number out of  $a_{n+1}, \dots, a_m, a_{m+1}, \dots, a_k, a'_{n+1}, a'_p, a''_{m+1}, \dots, a''_q$  is not equal to  $a$  then inequality  $d_a(u, w) \leq d_a(u, v) + d_a(v, w)$  is satisfied due



**Fig. 6** “Split-off” of three sequences.

to our assumption that  $\infty + \infty = \infty$  and  $\infty + r = \infty$  for each non-negative real number  $r$ . Otherwise,

$$\begin{aligned}
 d_a(u, w) &= \left( \sum_{i=n+1}^m d_i + \sum_{i=m+1}^k d_i \right) + \sum_{i=n+1}^p d'_i \\
 &\leq \left( \sum_{i=n+1}^m d_i + \sum_{i=m+1}^k d_i \right) + \sum_{i=n+1}^p d'_i + 2 \sum_{i=m+1}^q d''_i \\
 &= \left( \sum_{i=m+1}^k d_i + \sum_{i=m+1}^q d''_i \right) + \left( \sum_{i=n+1}^m d_i + \sum_{i=m+1}^q d''_i + \sum_{i=n+1}^p d'_i \right) \\
 &= d_a(u, v) + d_a(v, w).
 \end{aligned}$$

*Case II:* Now, assume that  $v$  splits off first. Then (see Figure 6), let

$$\begin{aligned}
 u &= s_0, (a_1, d_1), \dots, s_n, (a_{n+1}, d_{n+1}), \dots, (a_m, d_m), s_m, (a_{m+1}, d_{m+1}), \dots, s_k, \\
 v &= s_0, (a_1, d_1), \dots, s_n, (a'_{n+1}, d'_{n+1}), \dots, (a'_p, d'_p), s'_p, \\
 w &= s_0, (a_1, d_1), \dots, s_n, (a_{n+1}, d_{n+1}), \dots, (a_m, d_m), s_m, (a''_{m+1}, d''_{m+1}), \dots, s''_q.
 \end{aligned}$$

If at least one number out of  $a_{n+1}, \dots, a_m, a_{m+1}, \dots, a_k, a'_{n+1}, a'_p, a''_{m+1}, \dots, a''_q$  is not equal to  $a$  then inequality  $d_a(u, w) \leq d_a(u, v) + d_a(v, w)$  is satisfied due to our assumption that  $\infty + \infty = \infty$  and  $\infty + r = \infty$  for each non-negative

real number  $r$ . Otherwise, by Definition 12,

$$\begin{aligned}
d_a(u, w) &= \sum_{i=m+1}^k d_i + \sum_{i=m+1}^q d''_i \\
&= \sum_{i=m+1}^k d_i + \sum_{i=m+1}^q d''_i + 2 \sum_{i=n+1}^p d'_i + 2 \sum_{i=n+1}^m d_i \\
&= \left( \sum_{i=n+1}^m d_i + \sum_{i=m+1}^k d_i + \sum_{i=n+1}^p d'_i \right) \\
&\quad + \left( \sum_{i=n+1}^p d'_i + \sum_{i=n+1}^m d_i + \sum_{i=m+1}^q d''_i \right) \\
&= d_a(u, v) + d_a(v, w).
\end{aligned}$$

□

**Definition 13**  $\pi(p) = \{w \in W \mid p \in \text{end}(w)\}$  for any atomic proposition  $p$ .

This concludes the definition of the canonical confidence model  $(W, \{d_a\}_{a \in A}, \pi)$ . Remaining lemmas in this section describe various properties of the canonical model needed for the proof of completeness.

**Lemma 9** For any two worlds  $u, v \in W$ , if  $\Box_a^c \varphi \in \text{end}(u)$  and  $d_a(u, v) \leq c$ , then  $\varphi \in \text{end}(v)$ .

*Proof* Consider  $w = \text{lcp}(u, v)$ . By Part 2 of Lemma 7, assumption  $\Box_a^c \varphi \in \text{end}(u)$  implies that  $\Box_a^{c-d_a(u, w)} \varphi \in \text{end}(w)$ . Note that  $d_a(u, w) + d_a(w, v) = d_a(u, v) \leq c$  by Definition 12. Hence, by Part 1 of Lemma 7,

$$\Box_a^{c-d_a(u, w)-d_a(w, v)} \varphi \in \text{end}(v).$$

Thus,  $\Box_a^{c-d_a(u, v)} \varphi \in \text{end}(v)$  by again Definition 12. Thus,  $\text{end}(v) \vdash \varphi$  by Truth axiom. Therefore,  $\varphi \in \text{end}(v)$  due to the maximality of the set  $\text{end}(v)$ . □

**Lemma 10** For any epistemic world  $w \in W$ , if  $\Box_a^c \varphi \notin \text{end}(w)$ , then there is a world  $u \in W$  such that  $\varphi \notin \text{end}(u)$  and  $d_a(w, u) = c$ .

*Proof* If  $c = 0$ , then choose  $u$  to be world  $w$ . Suppose that  $\varphi \in \text{end}(w)$ . Thus,  $\text{end}(w) \vdash \Box_a^0 \varphi$  by Zero Confidence axiom. Hence,  $\Box_a^0 \varphi \in \text{end}(w)$  due to the maximality of set  $w$ , which contradicts the assumption of the lemma.

Suppose that  $c > 0$ . We first show that the following set is consistent

$$X = \{\neg \varphi\} \cup \{\Box_a^d \psi \mid \Box_a^{d+c} \psi \in \text{end}(w)\} \cup \{\neg \Box_a^{d+c} \chi \mid \neg \Box_a^d \chi \in \text{end}(w)\}.$$

Assume the opposite. Thus, there must exist

$$\Box_a^{d_1+c} \psi_1, \dots, \Box_a^{d_n+c} \psi_n, \neg \Box_a^{d'_1} \chi_1, \dots, \neg \Box_a^{d'_m} \chi_m \in \text{end}(w) \quad (1)$$

such that

$$\Box_a^{d_1} \psi_1, \dots, \Box_a^{d_n} \psi_n, \neg \Box_a^{d_1+c} \chi_1, \dots, \neg \Box_a^{d_m+c} \chi_m \vdash \varphi.$$

Hence, by Deduction theorem for propositional logic,

$$\vdash \Box_a^{d_1} \psi_1 \rightarrow (\dots (\Box_a^{d_n} \psi_n \rightarrow (\neg \Box_a^{d_1+c} \chi_1 \rightarrow (\dots (\neg \Box_a^{d_m+c} \chi_m \rightarrow \varphi) \dots))) \dots).$$

Thus, by Necessitation rule,

$$\vdash \Box_a^c (\Box_a^{d_1} \psi_1 \rightarrow (\dots (\Box_a^{d_n} \psi_n \rightarrow (\neg \Box_a^{d_1+c} \chi_1 \rightarrow (\dots (\neg \Box_a^{d_m+c} \chi_m \rightarrow \varphi) \dots))) \dots)).$$

By Distributivity axiom and Modus Ponens rule,

$$\Box_a^c \Box_a^{d_1} \psi_1 \vdash \Box_a^c (\Box_a^{d_2} \psi_2 \rightarrow (\dots \rightarrow (\neg \Box_a^{d_m+c} \chi_m \rightarrow \varphi) \dots)).$$

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

$$\Box_a^c \Box_a^{d_1} \psi_1, \dots, \Box_a^c \Box_a^{d_n} \psi_n, \Box_a^c \neg \Box_a^{d_1+c} \chi_1, \dots, \Box_a^c \neg \Box_a^{d_m+c} \chi_m \vdash \Box_a^c \varphi.$$

By Positive Introspection axiom applied  $n$  times,

$$\Box_a^{d_1+c} \psi_1, \dots, \Box_a^{d_n+c} \psi_n, \Box_a^c \neg \Box_a^{d_1+c} \chi_1, \dots, \Box_a^c \neg \Box_a^{d_m+c} \chi_m \vdash \Box_a^c \varphi.$$

By Negative Introspection axiom applied  $m$  times,

$$\Box_a^{d_1+c} \psi_1, \dots, \Box_a^{d_n+c} \psi_n, \neg \Box_a^{d_1} \chi_1, \dots, \neg \Box_a^{d_m} \chi_m \vdash \Box_a^c \varphi.$$

Hence,  $end(w) \vdash \Box_a^c \varphi$ , due to assumption (1). Thus, due to the maximality of  $end(w)$ , we have  $\Box_a^c \varphi \in end(w)$ , which contradicts the assumption of the lemma. Therefore, set  $X$  is consistent. Let  $\hat{X}$  be a maximal consistent extension of  $X$ . Define  $u$  to be the extension of sequence  $w$  such that  $u = w, (a, c), \hat{X}$ . By Definition 8 and due to the assumption  $c > 0$ , we have  $u \in W$ . By Definition 11,  $d(w, u) = c$ .  $\square$

**Lemma 11**  $w \Vdash \varphi$  if and only if  $\varphi \in end(w)$ , for any epistemic world  $w \in W$  and any formula  $\varphi \in \Phi$ .

*Proof* We prove the lemma by induction on the structural complexity of formula  $\varphi$ . If  $\varphi$  is an atomic proposition, then the required follows from Definition 13 and Definition 7. If formula  $\varphi$  is of the form  $\neg\psi$  or  $\psi \rightarrow \chi$ , then the required follows from maximality and consistency of the set  $end(w)$ , Definition 7 and the induction hypothesis in the standard way. Suppose that  $\varphi$  is of the form  $\Box_a^c \psi$ .

( $\Rightarrow$ ): If  $\Box_a^c \psi \notin end(w)$ , then, by Lemma 10, there exists  $u \in W$  such that  $d_a(w, u) = c$  and  $\psi \notin end(u)$ . Thus, by the induction hypothesis,  $u \not\Vdash \psi$ . Therefore, by Definition 7,  $w \not\Vdash \Box_a^c \psi$ .

( $\Leftarrow$ ): Assume that  $w \not\Vdash \Box_a^c \psi$ . Hence, by Definition 7, there must exist  $u \in W$  such that  $u \not\Vdash \psi$  and  $d_a(w, u) \leq c$ . Thus, by the induction hypothesis,  $\psi \notin end(u)$ . Therefore, by Lemma 9,  $\Box_a^c \psi \notin end(w)$ .  $\square$

### 5.3 Completeness Theorem

We are now ready to state and prove the completeness theorem for (not necessarily finite) confidence models.

**Theorem 1** *If  $w \Vdash \varphi$  for each epistemic world  $w$  of each confidence model, then  $\vdash \varphi$ .*

*Proof* Suppose that  $\not\vdash \varphi$ . Let  $s_0$  be any maximal consistent set containing  $\neg\varphi$ . Consider the canonical model defined, based on  $s_0$ , in Section 5.2. Note that single-element sequence  $s_0$  is a world of the canonical model. By Lemma 11,  $s_0 \not\vdash \varphi$ .  $\square$

## 6 Completeness for Finite Confidence Models

In this section we use the truncation operation on metric spaces to strengthen our completeness theorem from the class of all confidence models to the class of finite confidence models.

### 6.1 Truncation of Metric Spaces

**Definition 14** For any metric space  $(W, d)$ , any  $u, v \in W$ , and any positive (“threshold”) real number  $t$ , let truncated distance  $d|_t$  be defined as

$$d|_t(u, v) = \begin{cases} d(u, v), & \text{if } d(u, v) \leq t, \\ t, & \text{otherwise.} \end{cases}$$

**Lemma 12**  *$(X, d|_t)$  is a finite metric space for each metric space  $(X, d)$  and each positive real threshold value  $t$ .*

*Proof* Identity of Indiscernibles and Symmetry properties follow from Definition 14 and Definition 5. To prove Triangle Inequality, for any world  $w \in W$  we prove that  $d|_t(u, v) \leq d|_t(u, w) + d|_t(w, v)$ . If  $\max\{d(u, w), d(w, v)\} \geq t$ , then  $\max\{d|_t(u, w), d|_t(w, v)\} = t$ . Thus,

$$d|_t(u, v) \leq t = \max\{d|_t(u, w), d|_t(w, v)\} \leq d|_t(u, w) + d|_t(w, v).$$

Otherwise,  $\max\{d(u, w), d(w, v)\} < t$ . Thus,  $d|_t(u, w) + d|_t(w, v) = d(u, w) + d(w, v)$ . Hence, by Triangle Inequality for metric  $d$ ,

$$d|_t(u, v) \leq d(u, v) \leq d(u, w) + d(w, v) = d|_t(u, w) + d|_t(w, v).$$

$\square$

## 6.2 Truncation of Confidence Models

**Definition 15** For any confidence model  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$  and any real number  $t$ , let truncation of this model with threshold value  $t$  be triple  $(W, \{d_a \downarrow_t\}_{a \in \mathcal{A}}, \pi)$ .

The following lemma follows from Definition 15 and Lemma 12.

**Lemma 13** For any any confidence model  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$  and any threshold value  $t > 0$ , truncation  $(W, \{d_a \downarrow_t\}_{a \in \mathcal{A}}, \pi)$  is a finite confidence model.

Informally,  $\text{rank}(\varphi)$  is the largest modality superscript  $c$  that appears inside formula  $\varphi$ . Formally,  $\text{rank}(\varphi)$  is defined below.

**Definition 16** For any formula  $\varphi \in \Phi$ , the non-negative value  $\text{rank}(\varphi)$  is defined recursively as follows:

1.  $\text{rank}(p) = 0$  for each atomic proposition  $p$ ,
2.  $\text{rank}(\neg\psi) = \text{rank}(\psi)$ , for each  $\psi \in \Phi$ ,
3.  $\text{rank}(\psi \rightarrow \chi) = \max\{\text{rank}(\psi), \text{rank}(\chi)\}$ , for each  $\psi, \chi \in \Phi$ ,
4.  $\text{rank}(\Box_a^c \psi) = \max\{c, \text{rank}(\psi)\}$ , for each  $c \geq 0, a \in \mathcal{A}, \psi \in \Phi$ .

**Lemma 14** If  $\Vdash$  is satisfiability relation of a confidence model  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$  and  $\Vdash'$  is satisfiability relation of model  $(W, \{d_a \downarrow_t\}_{a \in \mathcal{A}}, \pi)$ , then  $w \Vdash \varphi$  if and only if  $w \Vdash' \varphi$ , for any  $\varphi \in \Phi$  such that  $\text{rank}(\varphi) \leq t$ .

*Proof* We prove by induction on structural complexity of formula  $\varphi$ . If  $\varphi$  is an atomic proposition  $p$ , then, by Definition 4, both  $w \Vdash p$  and  $w \Vdash' p$  are equivalent to statement  $w \in \pi(p)$ . Thus,  $w \Vdash p$  if and only if  $w \Vdash' p$ . Cases when  $\varphi$  is a negation or an implication follow from the induction hypothesis, Definition 4, and Definition 7. Assume now that  $\varphi$  is  $\Box_a^c \psi$ . Then, by Definition 16,  $c \leq t$  and  $\text{rank}(\psi) \leq t$ .

( $\Rightarrow$ ) : Suppose that  $w \not\Vdash' \Box_a^c \psi$ . Thus, by Definition 4, there is  $u \in W$  such that  $d_a \downarrow_t(w, u) \leq c$  and  $u \not\Vdash' \psi$ . Then,  $d_a(w, u) \leq c$  due to  $c \leq t$  and Definition 14. Also, by the induction hypothesis,  $u \not\Vdash \psi$ . Therefore,  $w \not\Vdash \Box_a^c \psi$  by Definition 7.  
 ( $\Leftarrow$ ) : Assume that  $w \not\Vdash \Box_a^c \psi$ . Thus, by Definition 7, there is  $u \in W$  such that  $d_a(w, u) \leq c$  and  $u \not\Vdash \psi$ . Then,  $d_a \downarrow_t(w, u) \leq d_a(w, u) \leq c$ . Also, by the induction hypothesis,  $u \not\Vdash' \psi$ . Therefore,  $w \not\Vdash' \Box_a^c \psi$  by Definition 4.  $\square$

## 6.3 Completeness Theorem

We now state and prove the main completeness result of this article.

**Theorem 2** If  $w \Vdash \varphi$  for each epistemic world  $w$  of each finite confidence model, then  $\vdash \varphi$ .

*Proof* Suppose that  $\not\vdash \varphi$ . By Theorem 1, there is an epistemic world  $w$  of a confidence model  $(W, \{d_a\}_{a \in \mathcal{A}}, \pi)$  such that  $w \not\Vdash \varphi$ . Let  $\Vdash'$  be satisfiability relation for truncated confidence model  $(W, \{d_a \downarrow_{\text{rank}(\varphi)}\}_{a \in \mathcal{A}}, \pi)$ . By Lemma 14,  $w \not\Vdash' \varphi$ .  $\square$

## 7 Discussion

In this article we studied epistemic modality expressing confidence of a single agent in a multiagent system. Similar to the notion of distributed knowledge modality in epistemic logic, one may consider using formula  $\Box_A^c \varphi$ , where  $A$  is an arbitrary set of agents, to express “group confidence” in statement  $\varphi$ . There are, however, at least two different ways to interpret this modality. One can say that formula  $\Box_A^c \varphi$  is satisfied in epistemic world  $w$  if formula  $\varphi$  is satisfied in each epistemic world that is no further than  $c$  away from  $w$  in the metric of each agent  $a \in A$ . Alternatively, one can require the total sum of all distances for all  $a \in A$  to be no more than  $c$ . These two definitions of group confidence have different modal properties. For example, the first definition satisfies property  $\Box_A^c \varphi \rightarrow (\Box_B^d \varphi \rightarrow \Box_{A \cup B}^{\max\{c,d\}} \varphi)$ . The second definition yields a stronger form of this property:  $\Box_A^c \varphi \rightarrow (\Box_B^d \varphi \rightarrow \Box_{A \cup B}^{c+d} \varphi)$ .

One can also define “common confidence” modality  $C_A^c \varphi$  among a group of agents  $A$ . Namely,  $w \Vdash C_A^c \varphi$  is true if for any sequence of epistemic worlds  $w_0, \dots, w_n$  and any  $a_1, \dots, a_n \in A$  if  $w = w_0$  and  $\sum_{1 \leq i \leq n} d_{a_i}(w_{i-1}, w_i) < c$ , then  $w_n \Vdash \varphi$ .

If Identity of Indiscernibles in Definition 5 is replaced with a weaker condition:  $d(u, u) = 0$ , then the resulting mathematical structure is known as *pseudo metric space*. Accordingly, if  $d_a$  in Definition 6 is assumed to be only a pseudo metric, then we call the resulting structure *pseudo confidence model*. Of a particular interest are pseudo confidence models in which any distance is either 0 or  $\infty$ . All epistemic worlds in such models are partitioned into classes such that distance between worlds in the same class is zero. Thus, any such model is isomorphic to a standard S5 Kripke model, where confidence modality  $\Box_a^0$  is translated to epistemic modality  $\Box_a$ .

There is a connection between confidence reasoning and judgement aggregation. Indeed, consider an epistemic world  $w \in W$  of an S5 Kripke model  $(W, \{\sim_a\}_{a \in A}, \pi)$ , where  $A$  is a set of agents. One can introduce judgement aggregation modality  $\Box^n \varphi$  to represent the fact that at most  $n$  agents in set  $A$  do not know that  $\varphi$  is true:

$$w \Vdash \Box^n \varphi \text{ if there is } A_0 \subseteq A \text{ such that (i) } |A \setminus A_0| \leq n \text{ and (ii) for any } a \in A_0, \text{ if } w \sim_a v, \text{ then } v \Vdash \varphi.$$

Alternatively, one can define judgement aggregation modality  $\Box^n \varphi$  to represent the fact that  $\varphi$  is true in all epistemic worlds indistinguishable from the current world by all except for at most  $n$  agents in set  $A$ :

$$w \Vdash \Box^n \varphi \text{ if } v \Vdash \varphi \text{ for all } v \text{ such that } |A \setminus \{a \in A \mid w \sim_a v\}| \leq n.$$

Note that function  $d(w, u) = |A \setminus \{a \in A \mid w \sim_a v\}|$  is a pseudo metric on the set of epistemic worlds and the alternative judgement aggregation modality is just confidence modality in the pseudo confidence model defined by this pseudo metric.

## References

1. Sergei Artemov. The logic of justification. *The Review of Symbolic Logic*, 1(04):477–513, 2008.
2. Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33. American Mathematical Society Providence, 2001.
3. Tristan Charrier, Florent Ouchet, and François Schwarzentruber. Big brother logic: reasoning about agents equipped with surveillance cameras in the plane. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 1633–1634. International Foundation for Autonomous Agents and Multiagent Systems, 2014.
4. Joseph Y Halpern and Yoav Shoham. A propositional modal logic of time intervals. *Journal of the ACM (JACM)*, 38(4):935–962, 1991.
5. Richard W Hamming. Error detecting and error correcting codes. *Bell System technical journal*, 29(2):147–160, 1950.
6. Subhash Khot and Rishi Saket. Hardness of embedding metric spaces of equal size. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 218–227. Springer, 2007.
7. Eugene F Krause. *Taxicab geometry: An adventure in non-Euclidean geometry*. Courier Dover Publications, 2012.
8. Robert S. Milnikel. The logic of uncertain justifications. *Annals of Pure and Applied Logic*, 165(1):305 – 315, 2014. The Constructive in Logic and Applications.
9. Lawrence S Moss and Rohit Parikh. Topological reasoning and the logic of knowledge: preliminary report. In *Proceedings of the 4th conference on Theoretical aspects of reasoning about knowledge*, pages 95–105. Morgan Kaufmann Publishers Inc., 1992.
10. Walter Rudin. Principles of mathematical analysis (international series in pure & applied mathematics). 1976.
11. Henrik Sahlgqvist. 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.
12. Mikhail Sheremet, Frank Wolter, and Michael Zakharyashev. A modal logic framework for reasoning about comparative distances and topology. *Annals of Pure and Applied Logic*, 161(4):534 – 559, 2010.