

# Knowing How to Plan

Yanjun Li

College of Philosophy,  
Nankai University, Tianjin, China

Yanjing Wang\*

Department of Philosophy,  
Peking University, Beijing, China

Various *planning-based know-how logics* have been studied in the recent literature. In this paper, we use such a logic to do *know-how-based planning* via model checking. In particular, we can handle the *higher-order epistemic planning* involving know-how formulas as the goal, e.g., find a plan to make sure  $p$  such that the adversary does not know how to make  $p$  false in the future. We give a PTIME algorithm for the model checking problem over finite epistemic transition systems and axiomatize the logic under the assumption of perfect recall.

## 1 Introduction

Standard Epistemic Logic (EL) mainly studies reasoning patterns of *knowing that  $\varphi$* , despite early contributions by Hintikka on formulating other *know-wh* expressions such as *knowing who* and *why* using first-order and higher-order modal logic. In recent years, there is a resurgence of interest on epistemic logics of *know-wh* powered by the new techniques for fragments of first-order modal logic based on the so-called *bundle modalities* packing a quantifier and a normal epistemic modality together [26, 24, 21]. Within the varieties of logics of *know-wh*, the logics of know-how received the most attention in AI (cf. e.g., [25, 8, 20, 16]).

Besides the inspirations from philosophy and linguistics (e.g., [23]), the idea of *automated planning* in AI also plays an important role in the developments of various logics of know-how. The core idea is (roughly) to interpret *knowing how to achieve  $\varphi$*  as a *de re* statement of knowledge: “*there is a plan such that you know that this plan will ensure  $\varphi$* ”. Here, depending on the exact notion of *plans* and how much we want to “ensure” that the plan works, there can be different semantics based on ideas from conformant planning and contingent planning in AI. However, as shown by Li and Wang [16], there is a logic core which is independent from the exact notions of plans underlying all these notions of know-how. We can actually unify different planning-based approaches in a powerful framework.

In this paper, we show that the connection between planning and know-how is not merely one way in terms of *planning-based know-how*, it also makes perfect sense to do *know-how-based planning* which generalizes the notion of planning to incorporate know-how based goals to be explained in the next subsection. As observed in [16], the typical epistemic planning problem given explicit models can be viewed as model checking problems in our framework with the know-how modality  $K_h$  in the language. For example,  $K_{h_1}(K_2\varphi \wedge \neg K_3K_2\varphi)$  captures the epistemic planning problem for agent 1 to ensure that agent 2 knows that  $\varphi$  and keep it a secret from agent 3. Such epistemic planning can also be done in other epistemic approaches, for example, by using dynamic epistemic logic (cf. e.g., [3, 5, 4, 7]). However, we can do much more with the  $K_h$  modality in hand, as explained below.

---

\*Corresponding author. Yanjing Wang thanks the support of NSSF grant 19BZX135.

## 1.1 Higher-order epistemic planning

A distinct feature of modal logic is that we can bring some notions of the meta-language to the object-language. This is not merely formalizing the existing meta-language concepts more precisely since the object language can open new possibilities. Consider the following multi-agent epistemic language of know-that and know-how:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid Kh_i\varphi$$

Since we have the know-how operators  $Kh_i$  in the fully compositional logical language, the goal of planning can involve the (Boolean combinations and nestings of) know-how formulas as well, and we call such planning problems *higher-order epistemic planning*, i.e., *planning about planning*. By *higher-order* we do not mean higher-order logic or higher-order epistemic goals in terms of nested know-that formulas.

Although in the single-agent scenario,  $Kh_1Kh_1\varphi$  is equivalent to  $Kh_1\varphi$  under reasonable conditions as shown in [8, 16], it still makes sense to do higher-order planning. For example, the planning problem to achieve  $\varphi$  with the future control to turn it off is expressed by  $Kh_1(\varphi \wedge Kh_1\neg\varphi)$ , which is not reducible to a formula without the nesting of  $Kh_1$ .<sup>1</sup>

In a multi-agent setting, given that different agent may have different abilities, the importance of higher-order planning is self-evident. Actually, it is a characteristic human instinct to plan in such a higher-order way. A one-year-old baby girl may not know how to open a bottle but she knows how to use her parents' knowledge-how to achieve her goal.<sup>2</sup> As a more concrete example about academic collaborations, it often happens that both researchers do not know how to prove a theorem independently, but one knows how to show a critical lemma which can simplify the original problem and thus enable the other to prove the final theorem using her expertise about the simplified statement. In our language, the situation can be expressed as  $\neg Kh_1\varphi \wedge \neg Kh_2\varphi \wedge Kh_1Kh_2\varphi$ . Of course, it depends on the cooperation of agent 2 to finally ensure the goal  $\varphi$  of agent 1. The nesting of know-how can go arbitrarily deep and also be interactive such as  $Kh_1Kh_2Kh_1Kh_2\varphi$ . Such planning based on others' knowledge-how is at the core of the arts of leadership and management in general.

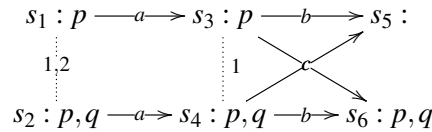
Higher-order planning also makes perfect sense in non-cooperative settings. For example, it is sometimes important to ensure not only the goal  $\varphi$  but also that the adversary cannot spoil it in the future. This is expressed by  $Kh_1(\varphi \wedge \neg Kh_2\neg\varphi)$  in our language of knowing how. It is also interesting to have mixed goals with both  $K$  and  $Kh$ , such as showing a commitment by achieving  $\varphi$  while letting the other know you will not be able to change it afterwards:  $Kh_1K_2(\varphi \wedge \neg Kh_1\neg\varphi)$ . In the more interactive game-like scenarios,  $Kh_1\neg Kh_2\neg Kh_1\varphi$  describe a winning strategy-like plan. We will come back to the related work using *Alternating-time Temporal Logic (ATL)*-like logics at the end of this introduction. Let us first see a concrete example.

**Example 1** *Suppose a patient is experiencing some rare symptom  $p$ . To know the cause ( $q$  or  $\neg q$ ), Doctor 1 in the local hospital suggests the patient first take a special fMRI scan ( $a$ ) which is only available in the local hospital at the moment, and bring the result to Doctor 2 in a hospital in another city, who is more experienced in examining the result from the scanner. Then Doctor 2*

<sup>1</sup>In contrast to normal modal logic,  $(Kh_i\varphi \wedge Kh_i\psi) \leftrightarrow Kh_i(\varphi \wedge \psi)$  is *invalid* thus  $Kh_iKh_i\varphi \leftrightarrow Kh_i\varphi$  cannot be applied, cf. e.g., [16].

<sup>2</sup>As new fathers, both authors know this well and sometimes it can be tiring to be a fully cooperative parent. At the same time, the parents are eager to know how to let their children acquire relevant knowledge-how as early as possible.

will then know the cause, depending on which different treatments ( $b$  or  $c$ ) should be performed by Doctor 2 to cure the patient. Intuitively, Doctor 1 knows how to let Doctor 2 know how to cure the patient although neither Doctor 1 nor Doctor 2 knows how to cure it without the help of the other. The situation is depicted below as a model with both epistemic relations (dotted lines labelled by 1,2) and action relations (solid lines labelled by  $a, b, c$ ). Note that only 1 can execute action  $a$  and only 2 can perform  $b$  or  $c$ . Reflexive epistemic arrows are omitted.



We would like to have  $\neg Kh_1 \neg p \wedge \neg Kh_2 \neg p \wedge Kh_1((K_2 q \vee K_2 \neg q) \wedge Kh_2 \neg p)$  hold at  $s_1$  according to the semantics to be introduced.

In this paper, to facilitate such higher-order planning using model checking, we extend the single-agent framework in [16] to a genuine multi-agent setting and study the model checking complexity in details. Technical and computational complications arise due to the introduction of multiple agents with different abilities and knowledge to base their own plans. To stay as close as the intended application of automated planning, we will restrict ourselves to finite models with perfect recall.<sup>3</sup> The model checking algorithm is based on a correspondence between execution trees in the models and the syntactically specified knowledge-based plans. It turns out that this correspondence result also enables us to significantly simplify the proofs of the soundness and completeness of the proof system compared with the method in [8, 16].

We summarize our contributions below.

- A multi-agent framework of logic of knowing how, which can handle the diversity of agents' abilities and formally specified plans based on each agent's own knowledge.
- A PTIME model checking algorithm for the full language.
- A complete axiomatization over finite models with perfect recall to ensure the semantics works as desired.<sup>4</sup>

## 1.2 Related work

Doing automated planning using model checking techniques is not a new idea (cf. e.g., [10]). The most distinctive feature of our framework, compared to other approaches to (epistemic) planning such as [5], is that we can express the planning problem in the object language and incorporate higher-order epistemic planning. Following [9], we base our framework on explicit-state models to lay out the theoretical foundation, instead of using more practical compact representations of models and actions which can generate the explicit-state models with certain properties (cf. [15, 16] for the discussions on the technical connections of the two). In contrast to the informal, model-dependent notions of plans such as policies, we use a plan specification language to formalize knowledge-based plans inspired by [13, 14, 27], but with more expressive

<sup>3</sup>cf. [16] for the discussions on the implicit assumption of perfect recall in epistemic planning and the technical difficulty of capturing it by axioms.

<sup>4</sup>Axiomatizations can also help us to do abstract syntactic reasoning about know-how without fixing a model.

branching conditions and more general constructors. Instead of the general structural operational semantics in [16], we give a more intuitive semantics for our knowledge-based plans in this paper.

Comparing to the coalition-based know-how approaches such as [20], we do not have any group know-how modality, but use a much more general notion of plans than the single-step joint action there. Besides the axiomatizations, which are the main focus in the previous work on the logic of know-how, we study the model checking complexity that is crucial for the application of our work to planning. A *second-order know-how modality*  $H_C^D$  was proposed in [19], where  $H_C^D \varphi$  says the coalition  $C$  knows how coalition  $D$  can ensure  $\varphi$  by a single-step joint action, though  $D$  may not know how themselves. Note that  $H_{\{i\}}^{\{j\}} \varphi$  is neither  $\text{Kh}_i \text{Kh}_j \varphi$  nor  $\text{K}_i \text{Kh}_j \varphi$  in our setting, where  $i$  may not know the plan that others may use to reach the goal.

The  $\text{Kh}_i$  operator is clearly also related to the strategy-based *de re* variants of the (single-agent) ATL operator  $\langle\langle i \rangle\rangle$  under imperfect information (cf. e.g., [22, 12]). In the setting of concurrent games of ATL, the strategies are functions from finite histories (or epistemic equivalence classes of them) to available actions, which induce the resulting plays of the given game that are usually infinite histories, on which temporal properties can be verified. This “global” notion of strategies leads to the problem of revocability of the strategies when nesting the  $\langle\langle i \rangle\rangle$  operators: can you change the strategy when evaluating  $\langle\langle i \rangle\rangle \varphi$  in the scope of another such operator [11, 1]? In contrast, it is crucial that our witness plans for know-how are not global and always terminate, and there is no controversy in the nested form of  $\text{Kh}_i$  operators in our setting. It is also crucial that the plans are executed sequentially in our setting, e.g.,  $\text{Kh}_i \text{Kh}_j \varphi$  simply means agent  $i$  knows how to make sure agent  $j$  knows how to make sure  $\varphi$  *after agent  $i$  finishes executing the witness plan for its know-how*. Moreover, while the *de re* variants of ATL are useful in reasoning about (imperfect information) concurrent games (cf. e.g., [18, 17]), our framework is closer to the standard automated planning in AI over transition systems. We leave it to the full version of the paper for a detailed technical comparison.

**Structure of the paper** In the rest of the paper, we first introduce briefly an epistemic language in Section 2 to be used to specify the branching conditions of our plan specification language in Section 3. In Section 4 we introduce the semantics of our know-how language and give a PTIME model checking algorithm. Finally, we obtain a complete axiomatization in Section 5 and conclude with future directions in Section 6.

Due to the strict space limitation, we have to omit some proofs.

## 2 Preliminaries

To specify the knowledge-based plans, we need the following formal language EAL to express the knowledge conditions.

**Definition 2 (EAL Language)** Let  $\mathbf{P}$  be a set of propositional letters,  $\mathbf{I}$  be a set of agents, and  $\mathbf{A}$  be a set of atomic actions. The language of Epistemic Action Logic ( $\text{EAL}_{\mathbf{I}}^{\mathbf{A}}$ ) is defined as follows:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid (\varphi \wedge \psi) \mid \text{K}_i \varphi \mid [a] \varphi$$

where  $p \in \mathbf{P}$ ,  $i \in \mathbf{I}$ , and  $a \in \mathbf{A}$ .

The auxiliary connectives and modalities  $\rightarrow, \vee, \langle a \rangle$  are defined as abbreviations as usual.

The semantics of EAL formulas is defined on finite models with both epistemic relations and transitions labelled by atomic actions, which will also be used for the know-how logic. In contrast with the single-agent model of [16], for each agent  $i$  we have a special set of actions  $\mathbf{A}_i$  that are executable by  $i$ . Note that an action may be executable by multiple agents. Moreover, we implicitly assume each agent is aware of all the actions, even for those not executable by itself.

**Definition 3 (Model)** An Epistemic Transition System (i.e., a model)  $\mathcal{M}$  is a tuple

$$\langle W, \{\sim_i \mid i \in \mathbf{I}\}, \{\mathbf{A}_i \mid i \in \mathbf{I}\}, \{Q(a) \mid a \in \bigcup_{i \in \mathbf{I}} \mathbf{A}_i\}, V \rangle$$

where:  $W$  is a non-empty set;  $\mathbf{A}_i$  is a set of actions for each  $i \in \mathbf{I}$ ;  $\sim_i \subseteq W \times W$  is an equivalence relation for each  $i \in \mathbf{I}$ ;  $Q(a) \subseteq W \times W$  is a binary relation for each  $a \in \bigcup_{i \in \mathbf{I}} \mathbf{A}_i$ ;  $V : W \rightarrow 2^{\mathbf{P}}$  is a valuation. A frame is a model without the valuation function.

**Notations** In the rest of the paper, we use  $\mathbf{A}$  to denote  $\bigcup_{i \in \mathbf{I}} \mathbf{A}_i$  for notational simplicity. We use  $[s]^i$  to denote the set  $\{t \in W \mid s \sim_i t\}$  and  $[W]^i$  to denote the set  $\{[s]^i \mid s \in W\}$ . We sometimes call the equivalence class  $[s]^i$  an  $i$ -belief-state. A model captures the agents' abilities to act and the uncertainty about the states.

$EAL_1^{\mathbf{A}}$  formulas are given truth values on pointed epistemic transition systems.

**Definition 4 (Semantics)** Given a pointed model  $(\mathcal{M}, s)$  where  $\mathcal{M} = \langle W, \{\sim_i \mid i \in \mathbf{I}\}, \{\mathbf{A}_i \mid i \in \mathbf{I}\}, \{Q(a) \mid a \in \mathbf{A}\}, V \rangle$  and a formula  $\varphi \in EAL_1^{\mathbf{A}}$ , the semantics is given below:

|  |                   |  |
|--|-------------------|--|
| $\mathcal{M}, s \models \top$                  | $\Leftrightarrow$ | always   |
| $\mathcal{M}, s \models p$                     | $\Leftrightarrow$ | $s \in \mathcal{V}(p)$   |
| $\mathcal{M}, s \models \neg \varphi$          | $\Leftrightarrow$ | $\mathcal{M}, s \not\models \varphi$                                     |
| $\mathcal{M}, s \models (\varphi \wedge \psi)$ | $\Leftrightarrow$ | $\mathcal{M}, s \models \varphi$ and $\mathcal{M}, s \models \psi$       |
| $\mathcal{M}, s \models K_i \varphi$           | $\Leftrightarrow$ | for all $t$ , if $s \sim_i t$ then $\mathcal{M}, t \models \varphi$      |
| $\mathcal{M}, s \models [a] \varphi$           | $\Leftrightarrow$ | for all $t$ , if $(s, t) \in Q(a)$ then $\mathcal{M}, t \models \varphi$ |

A formula is valid on a frame if it is true on all the pointed models based on that frame.

Let  $X$  be a set of states. We sometimes use  $\mathcal{M}, X \models \varphi$  to denote that  $\mathcal{M}, s \models \varphi$  for all  $s \in X$ . Thus  $\mathcal{M}, X \not\models \varphi$  denotes that  $\mathcal{M}, s \not\models \varphi$  for some  $s \in X$ , e.g., we may write  $\mathcal{M}, [s]^i \models K_i \varphi$  or  $\mathcal{M}, [s]^i \not\models K_i \varphi$ .

Here are some standard results about bisimulation that will be useful in the later technical discussion (cf. e.g., [2]).

**Definition 5 (Bisimulation)** Given a model  $\mathcal{M} = \langle W, \{\sim_i \mid i \in \mathbf{I}\}, \{\mathbf{A}_i \mid i \in \mathbf{I}\}, \{Q(a) \mid a \in \bigcup_{i \in \mathbf{I}} \mathbf{A}_i\}, V \rangle$ , a non-empty symmetric binary relation  $Z$  on the domain of  $\mathcal{M}$  is called a bisimulation iff whenever  $(w, v) \in Z$  the following conditions are satisfied:

- (1). For any  $p \in \mathbf{P} : p \in V(w) \iff p \in V(v)$ .
- (2). for any  $i \in \mathbf{I}$ , if  $w \sim_i w'$  for some  $w' \in W$  then there exists a  $v' \in W$  such that  $v \sim_i v'$  and  $w'Zv'$ .
- (3). for any  $a \in \bigcup_{i \in \mathbf{I}} \mathbf{A}_i$ , if  $(w, w') \in Q(a)$  for some  $w' \in W$  then there exists a  $v' \in W$  such that  $(v, v') \in Q(a)$  and  $w'Zv'$ .

The pointed models  $(\mathcal{M}, w)$  and  $(\mathcal{M}, v)$  are said to be bisimilar  $(\mathcal{M}, w \simeq \mathcal{M}, v)$  if there is a bisimulation  $Z$  such that  $(w, v) \in Z$ .

We use  $\mathcal{M}, [s]^i \Leftrightarrow \mathcal{M}, [s']^i$  to denote that for each  $t' \in [s']^i$ , there exists  $t \in [s]^i$  such that  $\mathcal{M}, t \Leftrightarrow \mathcal{M}, t'$  and vice versa. Given two pointed models  $(\mathcal{M}, s)$  and  $(\mathcal{M}, s')$ , we use  $\mathcal{M}, s \equiv_{EAL_{\mathbf{I}}^A} \mathcal{M}, s'$  to denote that for each  $\varphi \in EAL_{\mathbf{I}}^A$ ,  $\mathcal{M}, s \models \varphi$  if and only if  $\mathcal{M}, s' \models \varphi$ . Similarly,  $\mathcal{M}, s \equiv_{EAL_{\mathbf{I}}^A | K_i} \mathcal{M}, s'$  denotes that for each  $K_i \varphi \in EAL_{\mathbf{I}}^A$ ,  $\mathcal{M}, s \models K_i \varphi$  if and only if  $\mathcal{M}, s' \models K_i \varphi$ .

Please note that  $\mathcal{M}, s \Leftrightarrow \mathcal{M}, s'$  implies  $\mathcal{M}, [s]^i \Leftrightarrow \mathcal{M}, [s']^i$ , but the other way around does not work in general.

As an adaption of the Hennessy-Milner theorem for modal logic (cf. e.g., [2]), we have:

**Proposition 6** *Given a finite model  $\mathcal{M}$ , for each state  $s$  in  $\mathcal{M}$ , we have that  $\mathcal{M}, s \equiv_{EAL_{\mathbf{I}}^A} \mathcal{M}, s'$  iff  $\mathcal{M}, s \Leftrightarrow \mathcal{M}, s'$ .*

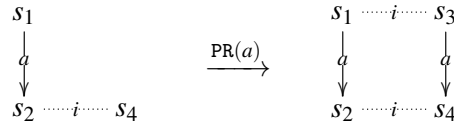
Moreover, if we only look at the  $K_i \varphi$  formulas, we have:

**Proposition 7** *Given a finite model  $\mathcal{M}$ , for each state  $s$  in  $\mathcal{M}$ , we have that  $\mathcal{M}, [s]^i \Leftrightarrow \mathcal{M}, [s']^i$  iff  $\mathcal{M}, s \equiv_{EAL_{\mathbf{I}}^A | K_i} \mathcal{M}, s'$ .*

Most approaches of epistemic planning implicitly assume the property of perfect recall (cf. [16, Sec. 6.5]), here we will also consider this extra property which will play a role in model checking.

**Definition 8 (Perfect recall)** *Given a model  $\mathcal{M}$ , for each  $i \in \mathbf{I}$  and each  $a \in \mathbf{A}_i$ , if  $(s_1, s_2) \in Q(a)$  and  $s_2 \sim_i s_4$  then there must exist a state  $s_3$  such that  $s_1 \sim_i s_3$  and  $(s_3, s_4) \in Q(a)$ .*

In words, if the agent cannot distinguish two states after doing  $a$  then it could not distinguish them before. Perfect recall (PR) corresponds to the following property depicted below:



### 3 A specification language for knowledge-based plans

Inspired by [13], we only consider knowledge-based plans for each agent. We introduce the following specification language which is a multi-agent variant of a fragment of the programming language introduced in [16].

**Definition 9 (Knowledge-based specification language  $KBP(i)$ )** *Based on  $EAL_{\mathbf{I}}^A$ ,  $KBP(i)$  is defined as follows:*

$$\pi ::= \varepsilon \mid a \mid (\pi; \pi) \mid \text{if } K_i \varphi \text{ then } \pi \text{ else } \pi$$

where  $\varepsilon$  is the empty plan,  $a \in \mathbf{A}_i$  and  $\varphi \in EAL_{\mathbf{I}}^A$ .

Note that the atomic action  $a$  in the above definition must belong to  $i$ , but the condition  $K_i \varphi$  can contain epistemic modalities and actions of other agents, e.g.,  $K_i \neg K_j [b] p$  is a legitimate condition for a knowledge-based plan for  $i$  even when  $b \notin \mathbf{A}_i$ .

We will abbreviate  $\text{if } K_i \varphi \text{ then } \pi_1 \text{ else } \pi_2$  as  $K_i \varphi ? \pi_1 : \pi_2$ . We use  $\pi^n$  to denote the program  $\underbrace{\pi; \dots; \pi}_n$ , and  $\pi^0$  is the empty program  $\varepsilon$ .

Let  $\mathcal{M}$  be an epistemic transition system,  $s$  be a state in  $\mathcal{M}$ , and  $\pi$  be a program in  $\text{KBP}(i)$ . The state set  $Q(\pi)(s)$ , which is the set of states on which executing  $\pi$  on  $s$  will terminate, is defined by induction on  $\pi$  as follows:

$$\begin{aligned} Q(\varepsilon)(s) &= \{s\} \\ Q(a)(s) &= \{t \mid (s, t) \in Q(a)\} \\ Q(\pi_1; \pi_2)(s) &= \{v \mid \text{there is } t \in Q(\pi_1)(s) : v \in Q(\pi_2)(t)\} \\ Q(\text{K}_i\varphi? \pi_1 : \pi_2)(s) &= \begin{cases} Q(\pi_1)(s) & \mathcal{M}, s \models \text{K}_i\varphi \\ Q(\pi_2)(s) & \mathcal{M}, s \not\models \text{K}_i\varphi \end{cases} \end{aligned}$$

We use  $Q(\pi)([s]^i)$  to denote the set  $\bigcup_{s' \in [s]^i} Q(\pi)(s')$ . Given the property of perfect recall, the resulting states of a knowledge-based plan have a nice property.

**Proposition 10** *If  $\mathcal{M}$  has perfect recall and  $\pi \in \text{KBP}(i)$ , then  $Q(\pi)([s]^i)$  is closed over  $\sim_i$ .*

Next we will define a notion of strong executability which is a generalized version of the one in [25].

**Definition 11 (Strong executability)** *We define a program  $\pi$  to be strongly executable on a pointed model  $(\mathcal{M}, s)$  by induction on  $\pi$  as follows: (1)  $\varepsilon$  is always strongly executable on  $\mathcal{M}, s$ ; (2)  $a$  is strongly executable on  $\mathcal{M}, s$  if  $Q(a)(s) \neq \emptyset$ ; (3)  $\pi_1; \pi_2$  is strongly executable on  $\mathcal{M}, s$  if  $\pi_1$  is strongly executable on  $\mathcal{M}, s$  and  $\pi_2$  is strongly executable on each  $t \in Q(\pi_1)(s)$ ; (4)  $\text{K}_i\varphi? \pi_1 : \pi_2$  is strongly executable on  $\mathcal{M}, s$  if either  $\mathcal{M}, s \models \text{K}_i\varphi$  and  $\pi_1$  is strongly executable on  $\mathcal{M}, s$ , or  $\mathcal{M}, s \not\models \text{K}_i\varphi$  and  $\pi_2$  is strongly executable on  $\mathcal{M}, s$ . We say that  $\pi$  is strongly executable on a set  $X$  of states if it is strongly executable on  $\mathcal{M}, s$  for each  $s \in X$ .*

To define the model checking algorithm later, we need to construct the tree of possible executions of a plan where each node is essentially an epistemic equivalence class (usually called a *belief state* in automated planning).

**Definition 12 (Execution tree)** *Given a model  $\mathcal{M}$  and an agent  $i \in \mathbf{I}$ , an execution tree of  $\mathcal{M}$  for  $i$  is a labeled tree  $\mathcal{T} = \langle N, E, L \rangle$ , where  $\langle N, E \rangle$  is a tree consisting of a nonempty set  $N$  of nodes and a set  $E$  of edges and  $L$  is a label function that labels each node in  $N$  with an  $i$ -belief-state and each edge in  $E$  with an action  $a \in \mathbf{A}_i$ , such that, for all  $(n, m) \in E$ , (1)  $L(n, m)$  is strongly executable on  $L(n)$ ; (2)  $L(n, m) = L(n', m')$  if  $(n', m') \in E$  and  $n = n'$ ; (3) there exists  $t \in L(m)$  such that  $t \in Q(L(n, m))(L(n))$ ; (4) for each  $t \in Q(L(n, m))(L(n))$ , there exists  $k \in N$  such that  $(n, k) \in E$  and  $t \in L(k)$ .*

We use  $r^{\mathcal{T}}$  (or simply  $r$ ) to refer to the root node of the execution tree  $\mathcal{T}$ . The following two propositions show the relations between execution trees and knowledge-based plans.

**Proposition 13** *Given a model  $\mathcal{M}$  and a program  $\pi \in \text{KBP}(i)$ , if  $\pi$  is strongly executable on  $[s]^i$ , then there is a finite execution tree  $\langle N, E, L \rangle$  of  $\mathcal{M}$  for  $i$  such that  $L(r) = [s]^i$  and that for each leaf node  $k$ ,  $L(k) \subseteq Q(\pi)([s]^i)$ .*

**Proposition 14** *Let  $\mathcal{T} = \langle N, E, L \rangle$  be a finite execution tree of  $\mathcal{M}$  for  $i$ . There exists  $\pi \in \text{KBP}(i)$  such that  $\pi$  is strongly executable on  $L(r)$  and that for each  $t \in Q(\pi)(L(r))$ , there exists a leaf node  $k$  such that  $\mathcal{M}, [t]^i \Leftrightarrow \mathcal{M}, L(k)$ .*

## 4 Model checking knowledge-how

Extending the work of [16], we introduce the multi-agent epistemic language of know-how and its semantics and study the model checking complexity in detail.

**Definition 15 (ELKh Language)** Let  $\mathbf{P}$  be a set of variables and  $\mathbf{I}$  be a set of agents, the language of epistemic logic with the know-how operator (ELKh) is defined as follows:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid \text{Kh}_i\varphi$$

where  $p \in \mathbf{P}$  and  $i \in \mathbf{I}$ .

**Definition 16 (Semantics)** The truth conditions for basic propositions, Boolean connectives and the epistemic operator  $K_i$  are as before in the case of EAL.

|   |
|---|
| $\mathcal{M}, s \models \text{Kh}_i\varphi \iff$ there is a plan $\pi \in \text{KBP}(i)$ such that <ol style="list-style-type: none"> <li>1. <math>\pi</math> is strongly executable on <math>[s]^i</math>;</li> <li>2. <math>\mathcal{M}, t \models \varphi</math> for each <math>t \in Q(\pi)([s]^i)</math>.</li> </ol> |
|---|

The readers can go back to Example 1.1 to verify the truth of the formula mentioned there.

We can show that ELKh is invariant under bisimulation defined in Definition 5 which also match labelled transitions (cf. [16]).

**Proposition 17** If  $\mathcal{M}, s \leftrightarrow \mathcal{M}, u$  then  $\mathcal{M}, s \models \varphi$  iff  $\mathcal{M}, u \models \varphi$  for all  $\varphi \in \text{ELKh}$ .

In the remainder of this section, we will propose a model checking algorithm for ELKh. The key part of the model checking algorithm is to check the  $\text{Kh}_i$ -formulas, and the problem of whether  $\mathcal{M}, s \models \text{Kh}_i\varphi$  depends on whether there is a good plan for  $\text{Kh}_i\varphi$ . Our strategy is to reduce the problem of whether there is a good plan for  $\text{Kh}_i\varphi$  to the nondeterministic planning problem in [6], which we will briefly introduce below.

**Definition 18 (Planning problem [6])** Let  $T = \langle D, \mathbf{A}, \{\overset{a}{\rightarrow} \mid a \in \mathbf{A}\} \rangle$  be a labeled transition system, where  $D$  is a finite set,  $\mathbf{A}$  is a set of actions, and  $\overset{a}{\rightarrow} \subseteq D \times D$  is a binary relation. A fully observable nondeterministic planning problem is a tuple  $P = \langle T, s, G \rangle$  where  $s \in D$  is the initial state, and  $G \subseteq D$  is the goal set.

Given a labeled transition system  $T$  (note that there is no epistemic relation), we often write  $(s, t) \in \overset{a}{\rightarrow}$  as  $s \overset{a}{\rightarrow} t$ . The sequence  $s_0 \overset{a_1}{\rightarrow} s_1 \overset{a_2}{\rightarrow} \dots s_n$  is called an execution trace of  $T$  from  $s_0$ . A partial function  $f : D \rightarrow \mathbf{A}$  is called a *policy*. An execution trace  $s_0 \overset{a_1}{\rightarrow} s_1 \overset{a_2}{\rightarrow} \dots s_n$  is induced by  $f$  if  $f(s_k) = a_{k+1}$  for all  $0 \leq k < n$ . It is complete w.r.t.  $f$  if  $s_n$  is not in the domain of  $f$ .

**Definition 19 (Strong plan)** A policy  $f$  is a strong plan for a planning problem  $P = \langle T, s, G \rangle$  if each complete execution trace induced by  $f$  from  $s$  is finite and terminates in  $G$ .

**Proposition 20** Let  $\mathcal{T} = \langle D, \mathbf{A}, \{\overset{a}{\rightarrow} \mid a \in \mathbf{A}\} \rangle$ . The existence of strong plans for a planning problem  $P = \langle T, s, G \rangle$  is in PTIME in the size of  $T$ .

**PROOF** It is shown in [6] that the procedure presented in Algorithm 1 always terminates and is correct for whether  $P$  has a strong plan. Moreover, the loop is executed at most  $|D| + 1$  times. Therefore, the procedure is in PTIME.  $\square$



**Algorithm 1:** A planning procedure

---

```

Procedure StrongPlanExistence ( $T, s, G$ ):
   $OldSA := \text{false}$ ;
   $SA := \emptyset$ ;
  while  $OldSA \neq SA$  and  $s \notin G \cup ST_s(SA)$  do           /*  $ST_s(SA)$  is the state set
     $\{v \in D \mid (v, a) \in SA \text{ for some } a \in \mathbf{A}\}$ . */
     $Prel := \{(v, a) \in D \times \mathbf{A} \mid v \notin G \cup ST_s(SA), v \text{ has } a\text{-successors, and all } a\text{-successors of}$ 
       $v \text{ are in } G \cup ST_s(SA)\}$ ;
     $OldSA := SA$ ;
     $SA := SA \cup Prel$ ;
  if  $s \in G \cup ST_s(SA)$  then return true;
  return false;

```

---

Now we define the nondeterministic planning problem that corresponds to  $\mathcal{M}, s \models \text{Kh}_i \varphi$ . Since the language of ELKh is a multi-agent one where different agents have different abilities, the corresponding nondeterministic planning problem must take the agent  $i$  into account. This makes the reduction more complex than that of single-agent know-how logics.

Given a model  $\mathcal{M} = \langle W, \{\sim_i \mid i \in \mathbf{I}\}, \{\mathbf{A}_i \mid i \in \mathbf{I}\}, \{Q(a) \mid a \in \mathbf{A}\}, V \rangle$  and an agent  $i$ , a planning problem for  $i$  is  $P(\mathcal{M}, i) = \langle T(\mathcal{M}, i), [w]^i, G \rangle$  where  $[w]^i$  is an  $i$ -belief-state,  $G$  is a set of  $i$ -belief-states, and  $T(\mathcal{M}, i) = \langle D, \mathbf{A}_i, \{\xrightarrow{a} \mid a \in \mathbf{A}_i\} \rangle$  is a labeled transition system where  $D = [W]^i$  and for each  $a \in \mathbf{A}_i$ ,  $[s]^i \xrightarrow{a} [t]^i \iff a$  is strongly executable on  $[s]^i$ , and there exists  $t' \in [t]^i : t' \in Q(a)([s]^i)$ .

We will show that the problem of whether  $\mathcal{M}, s \models \text{Kh}_i \varphi$  indeed can be reduced to a nondeterministic planning problem for  $i$  (Lemma 23). Before that, we need two auxiliary propositions.

**Proposition 21** *Given a model  $\mathcal{M}$  and an agent  $i$ , if a program  $\pi \in \text{KBP}(i)$  is strongly executable on  $[s]^i$ , then the planning problem  $P = \langle T(\mathcal{M}, i), [s]^i, G \rangle$  where  $G = \{[v]^i \mid \mathcal{M}, [v]^i \Leftrightarrow \mathcal{M}, [t]^i \text{ for some } t \in Q(\pi)([s]^i)\}$  has a strong plan.*

**Proposition 22** *Given  $\mathcal{M}$  and  $i$ , if a planning problem for  $i$ ,  $P(\mathcal{M}, i) = \langle T(\mathcal{M}, i), [s]^i, G \rangle$ , has a strong plan, then there exists  $\pi \in \text{KBP}(i)$  such that  $\pi$  is strongly executable on  $[s]^i$  and that for each  $t \in Q(a)([s]^i)$ , there exists  $[v]^i \in G$  such that  $\mathcal{M}, t \Leftrightarrow \mathcal{M}, v$ .*

**Lemma 23** *Given a model  $\mathcal{M}$ , we have that  $\mathcal{M}, s \models \text{Kh}_i \varphi$  iff the planning problem  $P(\mathcal{M}, i) = \langle T(\mathcal{M}, i), [s]^i, G \rangle$  where  $G = \{[t]^i \mid \mathcal{M}, [t]^i \models \text{K}_i \varphi\}$  has a strong plan.*

**PROOF** If  $\mathcal{M}, s \models \text{Kh}_i \varphi$ , it follows that there exists  $\pi \in \text{KBP}(i)$  such that  $\pi$  is strongly executable on  $[s]^i$  and that  $\mathcal{M}, t \models \varphi$  for each  $t \in Q(\pi)([s]^i)$ . By Proposition 10, it follows that  $\mathcal{M}, [t]^i \models \text{K}_i \varphi$  for each  $t \in Q(\pi)([s]^i)$ . Let  $G'$  be the set  $\{[v]^i \mid \mathcal{M}, [v]^i \Leftrightarrow \mathcal{M}, [t]^i \text{ for some } t \in Q(\pi)([s]^i)\}$ . By Proposition 17, we have that  $\mathcal{M}, [v]^i \models \text{K}_i \varphi$  for each  $[v]^i \in G'$ . It follows that  $G' \subseteq G$ . Let  $P'$  be the planning problem  $\langle T(\mathcal{M}, i), [s]^i, G' \rangle$ . By Proposition 21, the planning problem  $P'$  has a strong plan. Since  $G' \subseteq G$ , it follows that the planning problem  $P(\mathcal{M}, i)$  also has a strong plan.

If the planning problem  $P(\mathcal{M}, i)$  has a strong plan, by Proposition 22, it follows that there exists  $\pi \in \text{KBP}(i)$  such that  $\pi$  is strongly executable on  $[s]^i$  and that for each  $t' \in Q(a)([s]^i)$ , there exists  $[t]^i \in G$  such that  $\mathcal{M}, t' \Leftrightarrow \mathcal{M}, t$ . Since  $\mathcal{M}, t \models \varphi$  for each  $[t]^i \in G$ , by Proposition 17, it follows that  $\mathcal{M}, t' \models \varphi$  for each  $t' \in Q(a)([s]^i)$ . We then have that  $\mathcal{M}, s \models \text{Kh}_i \varphi$ .  $\square$

Finally, we are ready to show the upper bound of model checking ELKh.

**Algorithm 2:** A model checking procedure

---

```

Procedure MC( $\mathcal{M}, w, \varphi$ ): /* whether  $\mathcal{M}, w \models \varphi$  */
  switch  $\varphi$  do /* Boolean cases are omitted */
  | case  $K_i\varphi$  do
  | | foreach  $w' \in [w]^i$  do
  | | | if not MC( $\mathcal{M}, w', \varphi$ ) then return false ;
  | | | return true;
  | case  $Kh_i\varphi$  do
  | |  $G := \emptyset$ ;
  | | foreach  $[s]^i$  in  $T(\mathcal{M}, i)$  do
  | | | if MC( $\mathcal{M}, s, K_i\varphi$ ) then  $G := G \cup \{[s]^i\}$  ;
  | | | return StrongPlanExistence ( $T(\mathcal{M}, i), [w]^i, G$ );

```

---

**Theorem 24** *The model checking problem of ELKh is in PTIME in the size of models.*

**PROOF** Algorithm 2 presents the model checking algorithm MC for ELKh on epistemic transition systems with perfect recall. By Lemma 23, we call the procedure StrongPlanExistence presented in Algorithm 1 to check whether  $\mathcal{M}, w \models Kh_i\varphi$ . Proposition 20 shows that the procedure StrongPlanExistence can be computed in PTIME of the size of  $T(\mathcal{M}, i)$ . Since the size of  $T^{\mathcal{M}}$  is bounded by the size of  $\mathcal{M}$ , it follows that Algorithm 1 is in PTIME in the size of  $\mathcal{M}$ .  $\square$

## 5 Axiomatization

In this section, we axiomatize our logic of know-how over finite models with perfect recall. Note that we cannot apply the generic completeness result in [16] since our setting is not only multi-agent but also with an extra condition of perfect recall. It turns out the axiomatization is the multi-agent variant of the system in [8] which has an extra axiom AxKhtoKhK than the proof system in [16]. Note that although the proof system is similar, the semantics of the Kh in [8] is based on model-dependent functions from equivalence classes to actions instead of the syntactically specified plan in our work.

The proof system SLKHC is presented in Table 1.

The property of perfect recall is sufficient to guarantee the validity of AxKhtoKhK, but the axiom cannot characterize the property, as it is shown in [16]. Moreover, the composition axiom AxKhKh in SLKHC is a weaker version of the axiom  $Kh(\varphi \vee Kh\psi) \rightarrow Kh(\varphi \vee \psi)$  in [16].

Due to the axiom AxKhtoKhK, the stronger composition axiom in [16] can be derived in SLKHC based on AxKhKh.

**Proposition 25**  $\vdash Kh_i(\varphi \vee Kh_i\psi) \rightarrow Kh_i(\varphi \vee \psi)$ .

Next, we will show that the proof system SLKHC is sound and complete over finite models with perfect recall. By Propositions 13 and 14, it follows that there is a correspondence between execution trees and plans. With this correspondence result, we can significantly simplify the proofs of the soundness and completeness of the proof system SLKHC than [8] and [16]. More

**Axioms**

|           |   |
|-----------|---|
| TAUT      | all axioms of propositional logic                     |
| DISTK     | $K_i p \wedge K_i(p \rightarrow q) \rightarrow K_i q$ |
| T         | $K_i p \rightarrow p$                                 |
| 4         | $K_i p \rightarrow K_i K_i p$                         |
| 5         | $\neg K_i p \rightarrow K_i \neg K_i p$               |
| AxKtoKh   | $K_i p \rightarrow Kh_i p$                            |
| AxKhtoKhK | $Kh_i p \rightarrow Kh_i K_i p$                       |
| AxKhtoKKh | $Kh_i p \rightarrow K_i Kh_i p$                       |
| AxKhKh    | $Kh_i Kh_i p \rightarrow Kh_i p$                      |
| AxKhbot   | $\neg Kh_i \perp$                                     |

**Rules**

|        |  |      |                                       |
|--------|--|------|---------------------------------------|
| MP     | $\frac{\varphi, \varphi \rightarrow \psi}{\psi}$                             | NECK | $\frac{\varphi}{K_i \varphi}$         |
| MONOKh | $\frac{\varphi \xrightarrow{\psi} \psi}{Kh_i \varphi \rightarrow Kh_i \psi}$ | SUB  | $\frac{K_i \varphi}{\varphi[\psi/p]}$ |

Table 1: SLKHC

specifically, the validity of Axiom AxKhKh, which was highly non-trivial in [8], follows from the fact that two subsequent execution trees can be combined into one execution tree. (The detailed proof can be found in the following theorem of soundness.) For the completeness, the corresponding execution tree plays an important role in the proof of the truth lemma (Lemma 34).

**Theorem 26 (Soundness)** *The proof system SLKHC is sound over finite models with perfect recall.*

**PROOF** The validity of axioms T, 4, 5 is due to the standard semantics for  $K_i$ .

The axiom AxKtoKh says if  $p$  is known then you know how to achieve  $p$ . Its validity is guaranteed the empty program. The axiom AxKhtoKhK is valid due to Proposition 10. The axiom AxKhtoKKh is the positive introspection axiom for  $Kh_i$  whose validity is due to the semantics for  $Kh_i$ . The axiom AxKhbot says we cannot guarantee contradiction, which indirectly requires that the plan should terminate at someplace and is guaranteed by the condition (1) of the semantics for  $Kh_i$ .

Finally, we will show the validity of the axiom AxKhKh, that is,  $\mathcal{M}, s \models Kh_i \varphi$  if  $\mathcal{M}, s \models Kh_i Kh_i \varphi$ . If  $\mathcal{M}, s \models Kh_i Kh_i \varphi$ , it follows that there is a program  $\pi \in \text{KBP}(i)$  such that  $\pi$  is strongly executable on  $[s]^i$  and that  $\mathcal{M}, t \models Kh_i \varphi$  for each  $t \in Q(\pi)([s]^i)$ . By Proposition 13, there is a finite execution tree  $\mathcal{T}$  of  $\mathcal{M}$  for  $i$  such that  $L(r) = [s]^i$  and that for each leaf node  $k$ ,  $L(k) \subseteq Q(\pi)([s]^i)$ . It follows that  $\mathcal{M}, L(k) \models Kh_i \varphi$  for each leaf node  $k$ . We then have that for each leaf node  $k$ , there is a program  $\pi_k \in \text{KBP}(i)$  such that  $\pi_k$  is strongly executable on  $L(k)$  and that  $\mathcal{M}, v \models \varphi$  for all  $v \in Q(\pi_k)(L(k))$ . By Proposition 13 again, for each leaf  $k$  of  $\mathcal{T}$ , there is a finite execution tree  $\mathcal{T}_k$  of  $\mathcal{M}$  for  $i$  such that  $L^{\mathcal{T}_k}(r^{\mathcal{T}_k}) = L(k)$  and that for each leaf node  $l$  of  $\mathcal{T}_k$ ,  $L^{\mathcal{T}_k}(l) \subseteq Q(\pi_k)(L(k))$ . Moreover, since  $\mathcal{M}, v \models \varphi$  for all  $v \in Q(\pi_k)(L(k))$ , it follows that  $\mathcal{M}, L^{\mathcal{T}_k}(l) \models K_i \varphi$  for each leaf node  $l$  of  $\mathcal{T}_k$ . We then construct an execution tree  $\mathcal{T}'$  of  $\mathcal{M}$  for  $i$  by extending each leaf node  $k$  of  $\mathcal{T}$  with  $\mathcal{T}_k$ . Since  $\mathcal{T}$  and all  $\mathcal{T}_k$  are finite, it follows that  $\mathcal{T}'$  is finite. Moreover, we have that the root of  $\mathcal{T}'$  is labeled with  $[s]^i$  and  $\mathcal{M}, L^{\mathcal{T}'}(l) \models K_i \varphi$  for each leaf node  $l$  of  $\mathcal{T}'$ . By Proposition 14, there is a program  $\pi' \in \text{KBP}(i)$  such that  $\pi'$  is strongly executable on  $[s]^i$  and that for each  $v \in Q(\pi')([s]^i)$ , there exists a leaf node  $l$  of  $\mathcal{T}'$  such that  $\mathcal{M}, [v]^i \Leftrightarrow \mathcal{M}, L^{\mathcal{T}'}(l)$ . Since  $\mathcal{M}, L^{\mathcal{T}'}(l) \models K_i \varphi$  for each leaf node  $l$  of  $\mathcal{T}'$ , by Proposition 17, we have that  $\mathcal{M}, v \models K_i \varphi$  for all  $v \in Q(\pi')([s]^i)$ . Thus, we have that  $\mathcal{M}, s \models Kh_i \varphi$ .  $\square$

Next, we will show the completeness of  $\text{SLKHC}$ . The key is to show that every consistent formula is satisfiable. We will construct a canonical model that consists of all atoms which are maximal consistent sets with respect to a closure set of formulas and show that every consistent formula is satisfied in the canonical model.

Given a set of formulas  $\Delta$ , let:  $\Delta|_{K_i} = \{K_i\varphi \mid K_i\varphi \in \Delta\}$ ,  $\Delta|_{\neg K_i} = \{\neg K_i\varphi \mid \neg K_i\varphi \in \Delta\}$ ,  $\Delta|_{Kh_i} = \{Kh_i\varphi \mid Kh_i\varphi \in \Delta\}$ ,  $\Delta|_{\neg Kh_i} = \{\neg Kh_i\varphi \mid \neg Kh_i\varphi \in \Delta\}$ . Let  $\Phi$  be a subformula-closed set of  $\text{ELKh}$ -formulas that is finite.

**Definition 27** The closure  $cl(\Phi)$  is defined as:

$$\Phi \cup \{K_i\varphi, \neg K_i\varphi \mid \varphi \in \Phi, i \in \mathbf{I} \text{ occurs in } \Phi\}.$$

If  $\Phi$  is finite, so is  $cl(\Phi)$ . Next, we will use it to build a canonical model with respect to  $\Phi$ .

**Definition 28 (Atom)** We enumerate the formulas in  $cl(\Phi)$  by  $\{\psi_0, \dots, \psi_h\}$  where  $h \in \mathbb{N}$ . The formula set  $\Delta = \{Y_k \mid k \leq h\}$  is an atom of  $cl(\Phi)$  if

- $Y_k = \psi_k$  or  $Y_k = \neg\psi_k$  for each  $\psi_i \in cl(\Phi)$ ;
- $\Delta$  is consistent in  $\text{SLKHC}$ .

The following two propositions are similar with their single-agent versions in [8].

**Proposition 29** Let  $\Delta$  be an atom of  $cl(\Phi)$  and  $K_i\varphi \in cl(\Phi)$ . If  $K_i\varphi \notin \Delta$  then there exists  $\Delta'$  such that  $\Delta'|_{K_i} = \Delta|_{K_i}$  and  $\neg\varphi \in \Delta'$ .

**Proposition 30** Let  $\Delta$  and  $\Delta'$  be two atoms of  $cl(\Phi)$  such that  $\Delta|_{K_i} = \Delta'|_{K_i}$ . We have  $\Delta|_{Kh_i} = \Delta'|_{Kh_i}$ .

**Definition 31 (Canonical model)** Given a subformula-closed set  $\Phi$ , the canonical model  $\mathcal{M}^\Phi = \langle W, \{\sim_i \mid i \in \mathbf{I}\}, \{\mathbf{A}_i \mid i \in \mathbf{I}\}, \{Q(a) \mid a \in \bigcup_{i \in \mathbf{I}} \mathbf{A}_i\}, V \rangle$  is defined as:

- $W = \{\Delta \mid \Delta \text{ is an atom of } cl(\Phi)\}$ ;
- for each  $i \in \mathbf{I}$ ,  $\Delta \sim_i \Gamma \iff \Delta|_{K_i} = \Gamma|_{K_i}$ ;
- for each  $i \in \mathbf{I}$ ,  $\mathbf{A}_i = \{(i, \varphi) \mid Kh_i\varphi \in \Phi\}$ ;
- for each  $(i, \varphi) \in \bigcup \mathbf{A}_i$ , we have that  $(\Delta, \Gamma) \in Q(\varphi) \iff Kh_i\varphi \in \Delta$  and  $K_i\varphi \in \Gamma$ ;
- for each  $p \in \Phi$ ,  $p \in V(\Delta) \iff p \in \Delta$ .

**Proposition 32** If  $\Gamma \in Q(i, \varphi)(\Delta)$  and  $\Gamma' \in [\Gamma]^i$ , then  $\Gamma' \in Q(i, \varphi)([\Delta]^i)$ . In other words, the model  $\mathcal{M}^\Phi$  has perfect recall.

**PROOF** We only need to show that  $\Gamma' \in Q(i, \varphi)(\Delta)$ . Since  $\Gamma \in Q(i, \varphi)(\Delta)$ , it follows that  $Kh_i\varphi \in \Delta$  and  $K_i\varphi \in \Gamma$ . Since  $\Gamma' \in [\Gamma]^i$ , it follows that  $K_i\varphi \in \Gamma'$ . Thus, we have that  $\Gamma' \in Q(i, \varphi)(\Delta)$ .  $\square$

**Proposition 33** Let  $\Delta$  be a state in  $\mathcal{M}^\Phi$  and  $(i, \psi) \in \mathbf{A}_i$  be executable at  $\Delta$ . If  $Kh_i\varphi \in \Delta'$  for all  $\Delta' \in Q(i, \psi)(\Delta)$  then  $Kh_i\varphi \in \Delta$ .

**PROOF** First, we show that  $K_i\psi$  is not consistent with  $\neg Kh_i\varphi$ . Since  $(i, \psi) \in \mathbf{A}_i$  is executable at  $\Delta$ , it follows that  $Kh_i\psi \in \Delta$  and that there is some state in  $Q(i, \psi)(\Delta)$  that contains  $K_i\psi$ . Moreover, it is obvious that  $Kh_i\varphi \in cl(\Phi)$ . Assume that  $K_i\psi$  is consistent with  $\neg Kh_i\varphi$ . By a Lindenbaum argument, there exists an atom  $\Gamma$  of  $cl(\Phi)$  such that  $\{K_i\psi, \neg Kh_i\varphi\} \subseteq \Gamma$ . By the definition of  $\mathcal{M}^\Phi$ , it follows that  $(\Delta, \Gamma) \in Q(i, \psi)$ . Since we know that  $Kh_i\varphi \in \Delta'$  for all  $\Delta' \in Q(i, \psi)(\Delta)$ , it follows

that  $\text{Kh}_i\varphi \in \Gamma$ . It is contradictory with the fact that  $\Gamma$  is consistent. Thus,  $\text{K}_i\psi$  is not consistent with  $\neg\text{Kh}_i\varphi$ . Hence, we have that  $\vdash \text{K}_i\psi \rightarrow \text{Kh}_i\varphi$ .

Since  $\vdash \text{K}_i\psi \rightarrow \text{Kh}_i\varphi$ , it follows by Rule  $\text{MONOKh}$  and Axiom  $\text{AxKhtokhK}$  that  $\vdash \text{Kh}_i\psi \rightarrow \text{Kh}_i\text{Kh}_i\varphi$ . Moreover, it follows by Axiom  $\text{AxKhKh}$  that  $\vdash \text{Kh}_i\psi \rightarrow \text{Kh}_i\varphi$ . Since we have shown that  $\text{Kh}_i\psi \in \Delta$ , we have that  $\text{Kh}_i\varphi \in \Delta$ .  $\square$

**Lemma 34 (Truth Lemma)** *For each  $\varphi \in \text{cl}(\Phi)$ ,  $\mathcal{M}^\Phi, \Delta \models \varphi$  iff  $\varphi \in \Delta$ .*

**PROOF** We prove it by induction on  $\varphi$ . The atomic and Boolean cases are straightforward. The case of  $\text{K}_i\varphi$  can be proved by Proposition 29. Next, we only focus on the case of  $\text{Kh}_i\varphi$ .

**Right to Left:** If  $\text{Kh}_i\varphi \in \Delta$ , we will show  $\mathcal{M}^\Phi, \Delta \models \text{Kh}_i\varphi$ . We first show that  $\text{K}_i\varphi$  is consistent. If not, namely  $\vdash \text{K}_i\varphi \rightarrow \perp$ , it follows by Rule  $\text{MONOKh}$  that  $\vdash \text{Kh}_i\text{K}_i\varphi \rightarrow \text{Kh}_i\perp$ . It follows by Axiom  $\text{AxKhbot}$  that  $\vdash \text{Kh}_i\text{K}_i\varphi \rightarrow \perp$ . Since  $\text{Kh}_i\varphi \in \Delta$ , it follows by Axiom  $\text{AxKhtokhK}$  that  $\Delta \vdash \perp$ , which is in contradiction with the fact that  $\Delta$  is consistent. Therefore,  $\text{K}_i\varphi$  is consistent.

By Lindenbaum's Lemma, there exists an atom  $\Gamma$  such that  $\text{K}_i\varphi \in \Gamma$ . By the definition of  $\mathcal{M}^\Phi$ , it follows that  $(i, \varphi) \in \mathbf{A}_i$  and that  $(\Delta', \Gamma) \in \mathcal{Q}(i, \varphi)$  for each  $\Delta' \in [\Delta]^i$ . It means  $(i, \varphi)$  is strongly executable on  $[\Delta]^i$ . For each  $\Theta \in \mathcal{Q}(i, \varphi)([\Delta]^i)$ , by the definition of  $\mathcal{M}^\Phi$ , it follows that  $\text{K}_i\varphi \in \Theta$ . By Axiom  $\text{T}$ , it follows that  $\varphi \in \Theta$ . By IH, we then have that  $\mathcal{M}^\Phi, \Theta \models \varphi$  for each  $\Theta \in \mathcal{Q}(i, \varphi)([\Delta]^i)$ . Thus, we have that  $\mathcal{M}^\Phi, \Delta \models \text{Kh}_i\varphi$ .

**Left to Right:** Suppose  $\mathcal{M}^\Phi, \Delta \models \text{Kh}_i\varphi$ , we will show  $\text{Kh}_i\varphi \in \Delta$ . Since  $\mathcal{M}^\Phi, \Delta \models \text{Kh}_i\varphi$ , it follows that there exists  $\pi \in \text{KBP}(i)$  such that  $\pi$  is strongly executable on  $[\Delta]^i$  and that  $\mathcal{M}^\Phi, \Gamma \models \varphi$  for all  $\Gamma \in \mathcal{Q}(\pi)([\Delta]^i)$ . By IH, we have that  $\varphi \in \Gamma$  for all  $\Gamma \in \mathcal{Q}(\pi)([\Delta]^i)$ . Moreover, for each  $\Gamma \in \mathcal{Q}(\pi)([\Delta]^i)$ , if  $\Gamma' \in [\Gamma]^i$ , by Proposition 32, it follows that  $\Gamma' \in \mathcal{Q}(\pi)([\Delta]^i)$ , and we then have that  $\varphi \in \Gamma'$ . Moreover, with Proposition 29, it is easy to check that  $\text{K}_i\varphi \in \Gamma$  for all  $\Gamma \in \mathcal{Q}(\pi)([\Delta]^i)$ .

Since  $\pi$  is strongly executable on  $[\Delta]^i$ , by Proposition 13, there is a finite execution tree  $\mathcal{T}$  of  $\mathcal{M}^\Phi$  for  $i$  such that  $L(r) = [\Delta]^i$  and that  $L(k) \subseteq \mathcal{Q}(\pi)([\Delta]^i)$ . Since we have shown that  $\text{K}_i\varphi \in \Gamma$  for all  $\Gamma \in \mathcal{Q}(\pi)([\Delta]^i)$ , it follows that if  $k$  is a leaf node then  $\text{K}_i\varphi$  is in all states in  $L(k)$ . Next, we firstly show the following claim:

for each node  $k$  and each  $\Theta \in L(k) : \text{Kh}_i\varphi \in \Theta$ .

Please note that the execution tree  $\mathcal{T}$  is finite. We prove the claim above by induction on the height of nodes. If the height of  $k$  is 0, it means that  $k$  is a leaf node. It follows that  $\text{K}_i\varphi \in \Theta$  for each  $\Theta \in L(k)$ . By Axiom  $\text{AxKtokh}$ , it follows that  $\text{Kh}_i\varphi \in \Theta$  for each  $\Theta \in L(k)$ . With the IH that the claim holds for each node whose height is less than  $h$ , we will show that the claim holds for nodes whose height is  $h \geq 1$ . Given a node  $k$  whose height is  $h$ , since  $h \geq 1$ , it follows that there is a node  $k'$  such that  $(k, k')$  is an edge in  $\mathcal{T}$  and the height of  $k'$  is less than  $h$ , in other words,  $k'$  is a child node of  $k$ . Let  $L(k) = [\Theta]^i$  and  $L(k, k') = (i, \psi)$ . Since  $\mathcal{T}$  is an execution tree, it follows that  $(i, \psi)$  is strongly executable on  $[\Theta]^i$ , and then  $(i, \psi)$  is executable on  $\Theta$ . For each  $\Theta' \in \mathcal{Q}(i, \psi)(\Theta) \subseteq \mathcal{Q}(L(k, k'))(L(k))$ , by the definition of execution trees, it follows that there is a  $k$ -child  $k''$  such that  $\Theta' \in L(k'')$ . Since  $k''$  is a child of  $k$ , by IH, it follows that  $\text{Kh}_i\varphi \in \Theta'$ . Thus, we have shown that  $\text{Kh}_i\varphi \in \Theta'$  for each  $\Theta' \in \mathcal{Q}(i, \psi)(\Theta)$ . By Proposition 33, it follows that  $\text{Kh}_i\varphi \in \Theta$ . Thus, we have shown the claim. Since the root is labeled with  $[\Delta]^i$ , we then have that  $\text{Kh}_i\varphi \in \Delta$ .  $\square$

Let  $\varphi$  be a consistent formula. By Proposition 29, it follows that there is an atom  $\Delta$  of  $\text{cl}(\Phi)$  such that  $\varphi \in \Delta$ , where  $\Phi$  is the set of subformulas of  $\varphi$ . By Proposition 32 and Lemma 34, it follows that  $\varphi$  is satisfiable over finite models with perfect recall. The completeness then follows:

**Theorem 35 (Completeness)** *The proof system  $\text{SLKHIC}$  is complete over finite models with perfect recall.*

From the construction of the canonical model  $\mathcal{M}^\Phi$  in Definition 31, we can see that both the state set  $W$  and the action set  $A$  are bounded. This implies that the logic has a small model property. Moreover, by Theorem 24, the decidability then follows:

**Theorem 36** *The logic  $ELKh$  is decidable.*

## 6 Conclusion

In this paper, we propose the notion of higher-order epistemic planning by using an epistemic logic of knowing how. The planning problems can be encoded using model checking problems in the framework, which can be computed in PTIME in the size of the the model. We also axiomatize the logic over finite models with perfect recall.

As for future work, besides many theoretical questions about the knowing how logic as it is, we may extend its expressive power to capture conditional knowledge-how, which is very useful in reasoning about planning problems. For example, one may say I know how to achieve  $\varphi$  given  $\psi$ . Note that it is very different from  $\text{Kh}_i(\psi \rightarrow \varphi)$  or  $\text{K}_i\psi \rightarrow \text{Kh}_i\varphi$ . The important difference is that such conditional knowledge-how is global, compared to the current local semantics of  $\text{Kh}_i$ . This is similar to the binary global  $\text{Kh}_i$  operator introduced in [25]. It would be interesting to combine the two notations of know-how in the same framework. On the practical side, we can consider the model checking problems over compact representations of the actions and states using the techniques in [15, 16] connecting the explicit-state models with the compact representations.

## References

- [1] Thomas Ågotnes, Valentin Goranko & Wojciech Jamroga (2007): *Alternating-time temporal logics with irrevocable strategies*. In Dov Samet, editor: *Proceedings of (TARK 2007)*, pp. 15–24, doi:10.1145/1324249.1324256.
- [2] Patrick Blackburn, Maarten de Rijke & Yde Venema (2002): *Modal Logic*. Cambridge University Press, doi:10.1017/CBO9781107050884.
- [3] Thomas Bolander & Mikkel Birkegaard Andersen (2011): *Epistemic planning for single and multi-agent systems*. *Journal of Applied Non-Classical Logics* 21(1), pp. 9–34, doi:10.3166/jancl.21.9-34.
- [4] Thomas Bolander, Tristan Charrier, Sophie Pinchinat & François Schwarzenrüber (2020): *DEL-based epistemic planning: Decidability and complexity*. *Artificial Intelligence* 287, p. 103304, doi:10.1016/j.artint.2020.103304.
- [5] Thomas Bolander, Martin Holm Jensen & François Schwarzenrüber (2015): *Complexity Results in Epistemic Planning*. In: *Proceedings of IJCAI '15*, pp. 2791–2797. Available at <http://ijcai.org/Abstract/15/395>.
- [6] A. Cimatti, M. Pistore, M. Roveri & P. Traverso (2003): *Weak, Strong, and Strong Cyclic Planning via Symbolic Model Checking*. *Artificial Intelligence* 147(1–2), p. 35–84, doi:10.1016/S0004-3702(02)00374-0.
- [7] Thorsten Engesser, Robert Mattmüller, Bernhard Nebel & Michael Thielscher (2021): *Game description language and dynamic epistemic logic compared*. *Artificial Intelligence* 292, p. 103433, doi:10.1016/j.artint.2020.103433.

- [8] Raul Fervari, Andreas Herzig, Yanjun Li & Yanjing Wang (2017): *Strategically knowing how*. In: *Proceedings of IJCAI '17*, pp. 1031–1038, doi:10.24963/ijcai.2017/143.
- [9] Malik Ghallab, Dana Nau & Paolo Traverso (2004): *Automated planning: theory and practice*. Elsevier.
- [10] Fausto Giunchiglia & Paolo Traverso (2000): *Planning as Model Checking*, doi:10.1007/10720246\_1.
- [11] Andreas Herzig (2015): *Logics of knowledge and action: critical analysis and challenges*. *Autonomous Agents and Multi-Agent Systems* 29(5), pp. 719–753, doi:10.1007/s10458-014-9267-z.
- [12] Wojciech Jamroga & Thomas Ågotnes (2007): *Constructive knowledge: what agents can achieve under imperfect information*. *Journal of Applied Non-Classical Logics* 17(4), pp. 423–475, doi:10.3166/jancl.17.423-475.
- [13] Jérôme Lang & Bruno Zanuttini (2012): *Knowledge-Based Programs as Plans - The Complexity of Plan Verification*. In: *ECAI '12*, pp. 504–509, doi:10.3233/978-1-61499-098-7-504.
- [14] Jérôme Lang & Bruno Zanuttini (2013): *Knowledge-Based Programs as Plans : Succinctness and the Complexity of Plan Existence (Extended Abstract)*. In: *TARK '13*. Available at <https://arxiv.org/abs/1310.6429>.
- [15] Yanjun Li & Yanjing Wang (2019): *Multi-agent Knowing How via Multi-step Plans: A Dynamic Epistemic Planning Based Approach*. In: *Proceedings of LORI VII*, pp. 126–139, doi:10.1007/978-3-662-60292-8\_10.
- [16] Yanjun Li & Yanjing Wang (2021): *Planning-Based Knowing How: A Unified Approach*. *Artificial Intelligence* 296, p. 103487, doi:10.1016/j.artint.2021.103487.
- [17] Bastien Maubert, Aniello Murano, Sophie Pinchinat, François Schwarzenrüber & Silvia Stranieri (2020): *Dynamic Epistemic Logic Games with Epistemic Temporal Goals*. In: *Proceedings of ECAI 2020*, IOS Press, pp. 155–162, doi:10.3233/FAIA200088.
- [18] Bastien Maubert, Sophie Pinchinat, François Schwarzenrüber & Silvia Stranieri (2020): *Concurrent Games in Dynamic Epistemic Logic*. In Christian Bessiere, editor: *Proceedings of IJCAI 2020*, ijcai.org, pp. 1877–1883, doi:10.24963/ijcai.2020/260.
- [19] Pavel Naumov & Jia Tao (2018): *Second-Order Know-How Strategies*. In: *Proceedings of AAMAS '18*, pp. 390–398, doi:10.5555/3237383.3237444.
- [20] Pavel Naumov & Jia Tao (2018): *Together we know how to achieve: An epistemic logic of know-how*. *Artificial Intelligence* 262, pp. 279–300, doi:10.1016/j.artint.2018.06.007.
- [21] Anantha Padmanabha, R. Ramanujam & Yanjing Wang (2018): *Bundled fragments of first-order modal logic: (un)decidability*. In: *Proceedings of FSTTCS '18*, doi:10.4230/LIPICs.FSTTCS.2018.43.
- [22] Pierre-Yves Schobbens (2004): *Alternating-time logic with imperfect recall*. *Electron. Notes Theor. Comput. Sci.* 85(2), pp. 82–93, doi:10.1016/S1571-0661(05)82604-0.
- [23] Jason Stanley & Timothy Williamson (2001): *Knowing how*. *The Journal of Philosophy*, pp. 411–444, doi:10.2307/2678403.
- [24] Yanjing Wang (2017): *A New Modal Framework for Epistemic Logic*. In Jérôme Lang, editor: *Proceedings of TARK '17, EPTCS* 251, pp. 515–534, doi:10.4204/EPTCS.251.38.
- [25] Yanjing Wang (2018): *A logic of goal-directed knowing how*. *Synthese* 10, pp. 4419–4439, doi:10.1007/s11229-016-1272-0.
- [26] Yanjing Wang (2018): *Beyond knowing that: a new generation of epistemic logics*. In: *Jaakko Hintikka on knowledge and game theoretical semantics*, Springer, Cham, pp. 499–533, doi:10.1007/978-3-319-62864-6\_21.
- [27] Bruno Zanuttini, Jérôme Lang, Abdallah Saffidine & François Schwarzenrüber (2020): *Knowledge-based programs as succinct policies for partially observable domains*. *Artificial Intelligence* 288, p. 103365, doi:10.1016/j.artint.2020.103365.