Strategically Knowing How

Raul Fervari
University of Córdoba
CONICET

Andreas Herzig
University of Toulouse
IRIT-CNRS

Yanjun Li
University of Groningen

Yanjing Wang[*]
Peking University

Abstract

In this paper, we propose a single-agent logic of goal-directed knowing how extending the standard epistemic logic of knowing that with a new knowing how operator. The semantics of the new operator is based on the idea that knowing how to achieve φ means that there exists a (uniform) strategy such that the agent knows that it can make sure φ. We give an intuitive axiomatization of our logic and prove the soundness, completeness and decidability of the logic. The crucial axioms relating knowing that and knowing how illustrate our understanding of knowing how in this setting. This logic can be used in representing and reasoning about both knowledge-that and knowledge-how.

1 Introduction

Standard epistemic logic focuses on reasoning about propositional knowledge expressed by *knowing that* φ [Hintikka, 1962]. However, in natural language, various other knowledge expressions are also frequently used, such as *knowing what*, *knowing how*, *knowing why*, and so on.

In particular, *knowing how* receives much attention in both philosophy and AI. Epistemologists debate about whether knowledge-how is also propositional knowledge (cf. the survey [Fantl, 2008]), e.g., whether *knowing how to swim* can be rephrased using *knowing that*. In AI, it is crucial to let autonomous agents *know how* to fulfill certain goals in robotics, game playing, decision making, and multi-agent systems (cf. the surveys [Gochet, 2013; Ågotnes *et al.*, 2015]). In fact, a large body of AI planning can be viewed as finding algorithms to let the autonomous planner *know how* to achieve some propositional goals, i.e., to obtain goal-directed knowledge-how. Here, both propositional knowledge and knowledge-how matter, especially in the planning problems where initial uncertainty and non-deterministic actions are present. From a logician's point of view, it is interesting to see how *knowing how* interacts with *knowing that*, and how they differ in their reasoning patterns. A logic of *knowing how* also helps us to find a consistency notion regarding knowledge database with knowing how expressions.

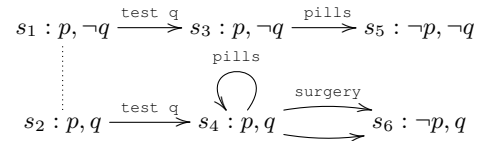[*]corresponding author (email: y.wang@pku.edu.cn)

Figure 1: A scenario representing how to cure the pain.

Example 1.1. *Consider the scenario where a doctor needs a plan to treat a patient and cure his pain (p), under the uncertainty about some possible allergy (q). If there is no allergy ($\neg q$) then simply taking some pills can cure the pain, and the surgery is not an option. On the other hand, in presence of the allergy, the pills may cure the pain or have no effect at all, while the surgery can cure the pain for sure. The model from Figure 1 represents this scenario with an additional action of testing whether q. The dotted line represents the initial uncertainty about q, and the test on q can eliminate this uncertainty (there is no dotted line between s_3 and s_4). According to the model, to cure the pain (guarantee $\neg p$) at the end, it makes sense to take the surgery if the result of the test of whether q is positive and take the pills otherwise. We can say the doctor in this case knows how to cure the pain.*

How to formalize the knowledge-how of the agent in such scenarios with uncertainty? Already since the early days of AI, people have been looking at it in the setting of logics of knowledge and action [McCarthy and Hayes, 1969; McCarthy, 1979; Moore, 1985; Lespérance *et al.*, 2000; van der Hoek *et al.*, 2000]. However, there has been no consensus on how to capture the logic of "knowing how" formally (cf. the recent surveys [Gochet, 2013] and [Ågotnes *et al.*, 2015]). The difficulties are well discussed in [Jamroga and Ågotnes, 2007] and [Herzig, 2015]: simply combining the existing modalities for "knowing that" and "ability" in a logical language like ATEL [van der Hoek and Wooldridge, 2003] does not lead to a genuine notion of "knowing how", e.g., knowing how to achieve p is not equivalent to merely knowing that there exists *some* strategy to make sure p. It does not work even when we replace the strategy by a *uniform strategy* where the agent has to choose the same action on indistinguishable states [Jamroga and Ågotnes, 2007]. Let $\varphi(x)$ express that x is a way to make sure some goal is achieved, and let $\mathcal{K}$ be the standard knowledge-that modality. There is a crucial distinction between the *de dicto* read-

ing of knowing how ($\mathcal{K}\exists x\varphi(x)$) and the desired *de re* reading ($\exists x\mathcal{K}\varphi(x)$) endorsed also by linguists and philosophers [Stanley and Williamson, 2001]. The latter intuitively implies the former, but not the other way round. For example, consider a variant of Example 1.1 where no test is available: then the doctor has *de dicto* knowledge-how to cure, but not *de re* knowledge. Proposals to capture the *de re* reading have been discussed in the literature, such as making the knowledge operator more constructive [Jamroga and Ågotnes, 2007], making the strategy explicitly specified [Herzig *et al.*, 2013; Belardinelli, 2014], or inserting $\mathcal{K}$ in-between an existential quantifier and the ability modality in see-to-it-that (STIT) logic [Broersen and Herzig, 2015].

In [Wang, 2015; 2016b], a new approach is proposed by introducing a single new modality $\mathcal{K}h$ of (conditional) goal-directed knowing how, instead of breaking it down into other modalities. This approach is in line with other *de re* treatments of non-standard epistemic logics of *knowing whether*, *knowing what* and so on (cf. [Wang, 2016a] for a survey). The semantics of $\mathcal{K}h$ is inspired by the idea of conformant planning based on linear plans [Smith and Weld, 1998; Yu *et al.*, 2016]. It is shown that $\mathcal{K}h$ is not a normal modality, e.g, knowing how to get drunk and knowing how to drive does not entail knowing how to drive when drunk. The work is generalized further in [Li and Wang, 2017; Li, 2017]. However, in these previous works, there was no explicit *knowing that* modality $\mathcal{K}$ in the language and the semantics of $\mathcal{K}h$ is based on linear plans, which does not capture the broader notion allowing branching plans or strategies that are essential in the scenarios like Example 1.1. In this paper, we extend this line of work in the following aspects:

- Both the *knowing how* modality $\mathcal{K}h$ and *knowing that* modality $\mathcal{K}$ are in the language, and we capture the interactions of the two explicitly by several axioms. Accordingly, we have both the action transitions and epistemic uncertainty in the models.

- In contrast to the essentially state-independent (global) semantics in [Wang, 2015; 2016b], we interpret $\mathcal{K}h$ locally w.r.t. the local uncertainty, in accordance with the standard semantics for the *knowing that* operator $\mathcal{K}$.

- Instead of linear plans in [Wang, 2015], the semantics of our $\mathcal{K}h$ operator is based on the existence of strategies (branching plans). In this way, the agent can make use of the knowledge learned during the execution of the plan, which helps us to capture situations where the agent intuitively knows how but does not have a linear plan.

The intuitive idea behind our semantics of $\mathcal{K}h\varphi$ is that the agent knows how to achieve φ iff (s)he has an executable *uniform strategy* σ such that the agent knows that:

- σ guarantees φ in the end given the uncertainty;
- σ always terminates after finitely many steps.

Note that for an agent to know how to make sure φ, it is not enough to find a plan which works *de facto*, but the agent should know it works in the end. This is a strong requirement inspired by planning under uncertainty, where the collection of final possible outcomes after executing the plan is required

to be a subset of the collection of the goal states [Geffner and Bonet, 2013].

Technically, our contributions are summarized as follows:

- A logical language with both $\mathcal{K}h$ and $\mathcal{K}$ operators with a semantics which fleshes out formally the above intuitions about knowing how.

- A complete axiomatization with intuitive axioms.

- Decidability of our logic.

The paper is organized as follows: Section 2 lays out the language and semantics of our framework; Section 3 proposes the axiomatization and proves its soundness; We prove the completeness of our proof system and show the decidability of the logic in Section 4 before we conclude with future work.

2 Language and Semantics

Let PROP be a countable set of propositional symbols.

Definition 2.1 (Language). *The language is defined by the following BNF where $p \in$ PROP:*

$$\varphi := p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathcal{K}\varphi \mid \mathcal{K}h\varphi.$$

We use $\bot, \vee, \rightarrow$ as usual abbreviations and write $\hat{\mathcal{K}}$ for $\neg\mathcal{K}\neg$.

Definition 2.2 (Models). *A model $\mathcal{M}$ is a quintuple $\langle W, ACT, \sim, \{\xrightarrow{a} \mid a \in ACT\}, V\rangle$ where:*

- *W is a non-empty set,*
- *ACT is a set of actions,*
- *$\sim \subseteq W \times W$ is an equivalence relation on W,*
- *$\xrightarrow{a} \subseteq W \times W$ is a binary relation for each $a \in ACT$*
- *$V : W \to 2^{PROP}$ is a valuation.*

Note that the labels in ACT do not appear in the language but only in the models. The graph in Example 1.1 represents a model with omitted self-loops of $\sim$ (dotted lines), and the equivalence classes induced by $\sim$ are $\{s_1, s_2\}, \{s_3\}, \{s_4\}, \{s_5\}, \{s_6\}$. In this paper we do not require any properties between $\sim$ and $\xrightarrow{a}$ to lay out the most general framework. We will come back to particular assumptions like *perfect recall* at the end of the paper. Given a model and a state s, if there exists t such that $s \xrightarrow{a} t$, we say that a is *executable* at s. Also note that the actions can be non-deterministic as in [Wang, 2015]. For each $s \in W$, we use $[s]$ to denote the equivalence class $\{t \in W \mid s \sim t\}$, and use $[W]$ to denote the collection of all the equivalence classes on W w.r.t. $\sim$. We use $[s] \xrightarrow{a} [t]$ to indicate that there are $s' \in [s]$ and $t' \in [t]$ such that $s' \xrightarrow{a} t'$. If there is $t \in W$ such that $[s] \xrightarrow{a} [t]$, we say a is executable at $[s]$.

Definition 2.3 (Strategies). *Given a model, a (uniformly executable) strategy is a partial function $\sigma : [W] \to ACT$ such that $\sigma([s])$ is executable at all $s' \in [s]$. Particularly, the empty function is also a strategy, the* empty *strategy.*

Note that the *executability* is as crucial as *uniformity*, without which knowledge-how may be trivialized. We use $\text{dom}(\sigma)$ to denote the domain of σ. Function σ can be seen as a binary relation such that $([s], a), ([s], b) \in \sigma$ implies $a = b$. Therefore, if τ is a restriction of σ, i.e. $\tau \subseteq \sigma$, it follows that $\text{dom}(\tau) \subseteq \text{dom}(\sigma)$, and $\tau([s]) = \sigma([s])$ for all $[s] \in \text{dom}(\tau)$.

Definition 2.4 (Executions)**.** *Given a strategy σ w.r.t a model $\mathcal{M}$, a* possible execution *of σ is a possibly infinite sequence of equivalence classes $\delta = [s_0][s_1] \cdots$ such that $[s_i] \xrightarrow{\sigma([s_i])} [s_{i+1}]$ for all $0 \le i < |\delta|$. Particularly, $[s]$ is a possible execution if $[s] \notin \text{dom}(\sigma)$. If the execution is a finite sequence $[s_0] \cdots [s_n]$, we call $[s_n]$ the* **leaf-node***, and $[s_i]$ $(0 \le i < n)$ an* **inner-node** *w.r.t. this execution. If it is infinite, then all $[s_i]$ $(i \in \mathbb{N})$ are* **inner-node***s. A possible execution of σ is* complete *if it is infinite or its* **leaf-node** *is not in $\text{dom}(\sigma)$.*

Given $\delta = [s_0] \cdots [s_n]$ and $\mu = [t_0] \cdots [t_m]$, we use $\delta \sqsubseteq \mu$ to denote that μ *extends* δ, i.e., $n \le m$ and $[s_i] = [t_i]$ for all $0 \le i \le n$. If $\delta \sqsubseteq \mu$, we define $\delta \sqcup \mu = \mu$. We use $\text{CELeaf}(\sigma, s)$ to denote the set of all **leaf-node**s of all the σ's complete executions (can be many due to non-determinism) starting from $[s]$, and $\text{CEInner}(\sigma, s)$ to denote the set of all the **inner-node**s of σ's complete executions starting from $[s]$. $\text{CELeaf}(\sigma, s) \cap \text{CEInner}(\sigma, s) = \emptyset$ since if $[s]$ is a **leaf-node** of a complete execution then σ is not defined at $[s]$.

Definition 2.5 (Semantics)**.** *Given a pointed model $\mathcal{M}, s$, the satisfaction relation $\vDash$ is defined as follows:*

$$\begin{aligned} \mathcal{M}, s \vDash p &\iff p \in \mathcal{V}(s) \\ \mathcal{M}, s \vDash \neg\varphi &\iff \mathcal{M}, s \nvDash \varphi \\ \mathcal{M}, s \vDash \varphi \wedge \psi &\iff \mathcal{M}, s \vDash \varphi \text{ and } \mathcal{M}, s \vDash \psi \\ \mathcal{M}, s \vDash \mathcal{K}\varphi &\iff \text{for all } s' : s \sim s' \text{ implies } \mathcal{M}, s' \vDash \varphi \\ \mathcal{M}, s \vDash \mathcal{K}h\varphi &\iff \text{there exists a strategy } \sigma \text{ such that} \end{aligned}$$
  1. $[t] \subseteq \llbracket \varphi \rrbracket$ *for all* $[t] \in \text{CELeaf}(\sigma, s)$
  2. *all its complete executions starting from $[s]$ are finite,*

where $\llbracket \varphi \rrbracket = \{ s \in W \mid \mathcal{M}, s \vDash \varphi \}$.

Note that the two conditions for σ in the semantics of $\mathcal{K}h$ reflect our two intuitions mentioned in the introduction. The implicit role of $\mathcal{K}$ in $\mathcal{K}h$ will become clearer when the axioms are presented. Going back to Example 1.1, we can verify that $\mathcal{K}h\neg p$ holds on s_1 and s_2 due to the strategy $\sigma = \{\{s_1, s_2\} \mapsto \texttt{test}, \{s_3\} \mapsto \texttt{pills}, \{s_4\} \mapsto \texttt{surgery}\}$. Note that $\text{CELeaf}(\sigma, s_1) = \{[s_5], [s_6]\} = \{\{s_5\}, \{s_6\}\}$ and $\llbracket \neg p \rrbracket = \{s_5, s_6\}$. On the other hand, $\mathcal{K}h\neg q$ is not true on s_1: although the agent can guarantee $\neg q$ *de facto* on s_1 by taking a strategy such that $\{s_1, s_2\} \mapsto \texttt{test}$ and $\{s_3\} \mapsto \texttt{pills}$, he cannot know it beforehand since nothing works at s_2 to make sure $\neg q$. Readers may also verify that $\mathcal{K}h(p \leftrightarrow q)$ holds at s_1 and s_2 (hint: a strategy is a partial function).

Let us compare our work to [Wang, 2015; 2016b], where the *knowing that* operator is not in the language. We include the knowledge operator $\mathcal{K}$ to precisely capture the interactions of *knowing that* and *knowing how* by the axioms to be introduced in the next section. Although the universal modality definable in the previous work can be viewed as a restricted kind of *knowing that* operator, it can only express global background knowledge to some extent. Moreover, in our models, both epistemic uncertainty and action transitions are present to model the changes of knowledge after actions, which is essential to facilitate strategies based on local knowledge, compared to the quite restricted linear plans of [Wang, 2015; 2016b]. Finally, the $\mathcal{K}h(\psi, \varphi)$ operator there ('the agent knows how to achieve φ given ψ') is state-independent. In the present perspective, the condition ψ there exactly defines an initial uncertainty set.

3 Axiomatization

3.1 The Proof System $\mathbb{SKH}$

Axioms

TAUT	all axioms of propositional logic
DISTK	$\mathcal{K}p \wedge \mathcal{K}(p \to q) \to \mathcal{K}q$
T	$\mathcal{K}p \to p$
4	$\mathcal{K}p \to \mathcal{K}\mathcal{K}p$
5	$\neg\mathcal{K}p \to \mathcal{K}\neg\mathcal{K}p$
AxKtoKh	$\mathcal{K}p \to \mathcal{K}hp$
AxKhtoKhK	$\mathcal{K}hp \to \mathcal{K}h\mathcal{K}p$
AxKhtoKKh	$\mathcal{K}hp \to \mathcal{K}\mathcal{K}hp$
AxKhKh	$\mathcal{K}h\mathcal{K}hp \to \mathcal{K}hp$
AxKhbot	$\mathcal{K}h\bot \to \bot$

Rules

MP	$\dfrac{\varphi, \varphi \to \psi}{\psi}$	NECK	$\dfrac{\varphi}{\mathcal{K}\varphi}$
MONOKh	$\dfrac{\varphi \to \psi}{\mathcal{K}h\varphi \to \mathcal{K}h\psi}$	SUB	$\dfrac{\varphi(p)}{\varphi[\psi/p]}$

Note that we have the $\mathbb{S}5$ axioms for $\mathcal{K}$. AxKtoKh says if p is known then you know how to achieve p by doing nothing (we allow the empty strategy). AxKhtoKhK reflects the first condition in the semantics that the goal is known after the executions. We will come back to this axiom at the end of the paper. Note that the termination condition is not fully expressible in our language, but AxKhbot captures part of it by ruling out strategies that have no terminating executions at all. AxKhKh essentially says that strategies can be composed. Its validity proof is quite involved and we devote the next subsection to it. Finally, AxKhtoKKh is the positive introspection axiom for $\mathcal{K}h$; it is valid due to uniformity of strategies. The corresponding negative introspection can be derived from axioms AxKhtoKKh, 5 and T:

Proposition 3.1. $\vdash \neg\mathcal{K}hp \to \mathcal{K}\neg\mathcal{K}hp$.

It is also not hard to show that in $\mathbb{SKH}$, $\mathcal{K}hp$ is provably equivalent to $\mathcal{K}\mathcal{K}hp$, $\mathcal{K}h\mathcal{K}p$, and $\mathcal{K}h\mathcal{K}hp$.

Note that we do not have the K axiom for $\mathcal{K}h$. Instead, we have the monotonicity rule MONOKh. In fact, the logic is not normal, as desired, e.g., $\mathcal{K}hp \wedge \mathcal{K}hq \to \mathcal{K}h(p \wedge q)$ is not valid: the existence of two different strategies for different goals does not imply the existence of a unified strategy to realize both goals.

3.2 Validity of AxKhKh

AxKhKh is about the "sequential" compositionality of strategies. Suppose on some pointed model there is a strategy σ to guarantee that we end up with the states where on each s of them we have some other strategy σ_s to make sure p ($\mathcal{K}h\mathcal{K}hp$). Since the strategies are uniform, we only need to consider some $\sigma_{[s]}$ for each $[s]$. Now to validate AxKhKh, we need to design a unified strategy to compose σ and those $\sigma_{[s]}$ into one strategy to still guarantee p ($\mathcal{K}hp$). The general idea is actually simple: first, order those leaf nodes $[s]$ (using Axiom of Choice); then by transfinite induction adjust $\sigma_{[s]}$ one by one to make sure these strategies can fit together as a unified strategy θ; finally, merge the relevant part of σ with θ into the desired strategy. We make this idea precise below. First we need an observation:

Proposition 3.2. *Given strategies τ and σ with $\tau \subseteq \sigma$, if $[s] \in \mathrm{dom}(\tau)$ and $\mathrm{dom}(\sigma) \cap \mathrm{CELeaf}(\tau, s) = \emptyset$, then a sequence is σ's complete execution from $[s]$ if and only if it is τ's complete execution from $[s]$.*

Proof. Left to Right: Let $[s_0] \cdots [s_n] \cdots$ be a σ's complete execution from $[s]$ (it also means that $[s] = [s_0]$, similarly in later proofs). We will show it is also a τ's complete execution from $[s]$. Firstly, we show it is a possible execution give τ from $[s]$. If it is not, there exists $[s_i]$ such that $[s_i]$ is not the leaf-node of this execution and such that $[s_i] \notin \mathrm{dom}(\tau)$. Let $[s_j]$ be the minimal equivalence class in the sequence with such properties. It follows that $[s_j] \in \mathrm{CELeaf}(\tau, s)$ and $[s_j] \in \mathrm{dom}(\sigma)$. These are contradictory with $\mathrm{dom}(\sigma) \cap \mathrm{CELeaf}(\tau, s) = \emptyset$.

Next we will show that $[s_0] \cdots [s_n] \cdots$ is a τ's complete execution from $[s]$. It is obvious if the sequence is infinite. If it is finite, let the leaf-node be $[s_m]$. It follows that $[s_m] \notin \mathrm{dom}(\sigma)$. Since $\tau \subseteq \sigma$, it follows $[s_m] \notin \mathrm{dom}(\tau)$. Therefore, the execution is complete given τ.

Right to Left: Let $[s_0] \cdots [s_n] \cdots$ be a τ's complete execution from $[s]$, we will show it is also a σ's complete execution from $[s]$. Since $\tau \subseteq \sigma$, it is also a possible execution given σ. If the execution is infinite, it is obvious. If it is finite, let the leaf-node be $[s_m]$. It follows that $[s_m] \in \mathrm{CELeaf}(\tau, s)$. Since $\mathrm{dom}(\sigma) \cap \mathrm{CELeaf}(\tau, s) = \emptyset$, it follows that $[s_m] \notin \mathrm{dom}(\sigma)$. Therefore, the execution is also complete give σ. □

Proposition 3.3. $\vDash \mathcal{K}h\mathcal{K}h\varphi \to \mathcal{K}h\varphi$.

Proof. Supposing $\mathcal{M}, s \vDash \mathcal{K}h\mathcal{K}h\varphi$, we will show that $\mathcal{M}, s \vDash \mathcal{K}h\varphi$. It follows by the semantics that there exists a strategy σ such that all σ's complete executions from $[s]$ are finite and $[t] \subseteq [\![\mathcal{K}h\varphi]\!]$ for all $[t] \in \mathrm{CELeaf}(\sigma, s)$ $(*)$. If $[s] \notin \mathrm{dom}(\sigma)$, then $\mathrm{CELeaf}(\sigma, s) = \{[s]\}$, and then it is trivial that $\mathcal{M}, s \vDash \mathcal{K}h\varphi$. Next we focus on the case of $[s] \in \mathrm{dom}(\sigma)$.

According to the well-ordering theorem (equivalent to Axiom of Choice), we assume $\mathrm{CELeaf}(\sigma, s) = \{S_i \mid i < \gamma\}$ where γ is an ordinal number and $\gamma \geq 1$. Let s_i be an element in S_i, then $[s_i] = S_i$. Since $\mathcal{M}, s_i \vDash \mathcal{K}h\varphi$ for each $i < \gamma$, it follows that for each $[s_i]$ there exists a strategy σ_i such that all σ_i's complete executions from $[s_i]$ are finite and $[v] \subseteq [\![\varphi]\!]$ for all $[v] \in \mathrm{CELeaf}(\sigma_i, s_i)$ ($\blacktriangleleft$). Next, in order to show $\mathcal{M}, s \vDash \mathcal{K}h\varphi$, we need to define a strategy τ. The definition consists of the following steps.

Step I. By induction on i, we will define a set of strategies τ_i where $0 \leq i < \gamma$. Let $f_i = \bigcup_{\beta < i} \tau_\beta$ and $D_i = \mathrm{CEInner}(\sigma_i, s_i) \setminus (\mathrm{dom}(f_i) \cup \{[v] \in \mathrm{CELeaf}(f_i, t) \mid [t] \in \mathrm{dom}(f_i)\})$; we define:

- $\tau_0 = \sigma_0|_{\mathrm{CEInner}(\sigma_0, s_0)}$;
- $\tau_i = f_i \cup (\sigma_i|_{D_i})$ for $i > 0$.

Claim 3.3.1. *We have the following results:*

1. *For each $0 \leq i < \gamma$, $\tau_j \subseteq \tau_i$ if $j < i$;*

2. *For each $0 \leq i < \gamma$, τ_i is a partial function;*

3. *For each $0 \leq i < \gamma$, $\mathrm{dom}(\tau_i) \cap \mathrm{CELeaf}(\tau_j, t) = \emptyset$ where $t \in \mathrm{dom}(\tau_j)$ if $j < i$;*

4. *For each $0 \leq i < \gamma$, if $\delta = [t_0][t_1] \cdots$ is a τ_i's complete execution from $[t] \in \mathrm{dom}(\tau_i)$ then $|\delta| = n$ for some $n \in \mathbb{N}$ and $[t_n] \subseteq [\![\varphi]\!]$;*

5. *For each $0 \leq i < \gamma$, $[s_i] \in \mathrm{dom}(\tau_i)$ or $[s_i] \subseteq [\![\varphi]\!]$.*

Proof of Claim 3.3.1:

1. It is obvious.

2. We prove it by induction on i. For the case of $i = 0$, it is obvious. For the case of $i = \alpha > 0$, it follows by the IH that τ_β is a partial function for each $\beta < \alpha$. Furthermore, it follows by (1) that $\tau_{\beta_1} \subseteq \tau_{\beta_2}$ for all $\beta_1 < \beta_2 < \alpha$. Thus, we have $f_\alpha = \bigcup_{\beta < \alpha} \tau_\beta$ is a partial function. Since σ_α is a partial function, in order to show τ_α is a partial function, we only need to show that $\mathrm{dom}(f_\alpha) \cap D_\alpha = \emptyset$. Since $D_\alpha = \mathrm{CEInner}(\sigma_\alpha, s_\alpha) \setminus \mathrm{dom}(f_\alpha) \setminus \{[v] \in \mathrm{CELeaf}(f_\alpha, t) \mid t \in \mathrm{dom}(f_\alpha)\}$, it is obvious that $\mathrm{dom}(f_\alpha) \cap D_\alpha = \emptyset$.

3. We prove it by induction on i. It is obvious for the case of $i = 0$. For the case of $i = \alpha > 0$, given $j < \alpha$ and $t \in \mathrm{dom}(\tau_j)$, we need to show that $\mathrm{dom}(\tau_\alpha) \cap \mathrm{CELeaf}(\tau_j, t) = \emptyset$. Supposing $[v] \in \mathrm{CELeaf}(\tau_j, t)$, we will show that $[v] \notin \mathrm{dom}(\tau_\alpha)$, namely $[v] \notin \mathrm{dom}(f_\alpha) \cup D_\alpha$. Since $j < \alpha$ and $f_\alpha = \bigcup_{\beta < \alpha} \tau_\beta$, it follows $t \in \mathrm{dom}(f_\alpha)$. Moreover, due to $D_\alpha = \mathrm{CEInner}(\sigma_\alpha, s_\alpha) \setminus (\mathrm{dom}(f_\alpha) \cup \{[v] \in \mathrm{CELeaf}(f_\alpha, t) \mid t \in \mathrm{dom}(f_\alpha)\})$, it follows $[v] \notin D_\alpha$.

Next, we only need to show $[v] \notin \mathrm{dom}(f_\alpha)$. Assuming $[v] \in \mathrm{dom}(f_\alpha)$, it follows that $[v] \in \mathrm{dom}(\tau_\beta)$ for some $\beta < \alpha$. There are two cases: $j < \beta$ or $j \geq \beta$. If $j < \beta$, it follows by the IH that $\mathrm{dom}(\tau_\beta) \cap \mathrm{CELeaf}(\tau_j, t) = \emptyset$. Contradiction. If $j \geq \beta$, it follows by (1) that $\tau_\beta \subseteq \tau_j$. Due to $[v] \in \mathrm{dom}(\tau_\beta)$, it follows $[v] \in \mathrm{dom}(\tau_j)$. It is contradictory with $[v] \in \mathrm{CELeaf}(\tau_j, t)$. Thus, we have $[v] \notin \mathrm{dom}(f_\alpha)$.

4. We prove it by induction on i. For the case of $i = 0$, due to $\mathrm{dom}(\tau_0) = \mathrm{CEInner}(\sigma_0, s_0)$, it follows that there is a σ_0's possible execution $[s_0] \cdots [s_m]$ such that $m \in \mathbb{N}$ and $[s_m] = [t]$. Let $\mu = [s_0] \cdots [s_{m-1}] \circ \delta$. (If $m = 0$ then $\mu = \delta$). Since δ is a τ_0's complete execution from $[t]$, it follows that μ is a σ_0's complete execution from $[s_0]$. It follows by ($\blacktriangleleft$) that μ is finite. Thus, $\delta = [t_0] \cdots [t_n]$ for some $n \in \mathbb{N}$. Since $[t_n] \in \mathrm{CELeaf}(\sigma_0, s_0)$, it follows by ($\blacktriangleleft$) that $[t_n] \subseteq [\![\varphi]\!]$.

For the case of $i = \alpha > 0$, there are two situations: $[t] \in \mathrm{dom}(f_\alpha)$ or $[t] \in D_\alpha$. If $[t] \in \mathrm{dom}(f_\alpha)$, it follows that $[t] \in \mathrm{dom}(\tau_\beta)$ for some $\beta < \alpha$. By (3), we have $\mathrm{dom}(\tau_\alpha) \cap \mathrm{CELeaf}(\tau_\beta, t) = \emptyset$. Since δ is a τ_α's complete execution, it follows by Proposition 3.2 that δ is also a τ_β's complete execution from $[t]$. It follows by the IH that $|\delta| = n$ for some $n \in \mathbb{N}$ and $[t_n] \subseteq [\![\varphi]\!]$.

If $[t] \in D_\alpha$, there are two cases: there exist $k < |\delta|$ and $\beta < \alpha$ s.t. $[t_k] \in \mathrm{dom}(\tau_\beta)$, or there do not exist such k and β. (Please note that $|\delta| > 1$ due to the fact that $\delta = [t_0] \cdots$ is τ_α's complete execution from $[t] \in \mathrm{dom}(\tau_\alpha)$).

- $[t_k] \in \mathrm{dom}(\tau_\beta)$ for some $k < |\delta|$ and some $\beta < \alpha$: It follows that $\mu = [t_k] \cdots$ is a τ_α's complete

execution from $[t_k]$. By (3) and Proposition 3.2, μ is a τ_β's complete execution from $[t_k]$. By IH, $\mu = [t_k] \cdots [t_{k+n}]$ for some $n \in \mathbb{N}$ and $[t_{k+n}] \subseteq [\![\varphi]\!]$. Therefore, $|\delta| = k + n$.

- If there does not exist $k < |\delta|$ and $\beta < \alpha$ s.t. $[t_k] \in \text{dom}(\tau_\beta)$, it follows that $\delta = [t_0] \cdots$ is a σ_α's possible execution from $[t]$. Since $[t] \in D_\alpha \subseteq \text{CEInner}(\sigma_\alpha, s_\alpha)$, there is a σ_α's possible execution $[s_0] \cdots [s_m]$ s.t. $m \in \mathbb{N}$, $[s_0] = [s_\alpha]$ and $[s_m] = [t]$. Let $\mu = [s_0] \cdots [s_{m-1}] \circ \delta$. (If $m = 0$ then $\mu = \delta$). It follows that μ is σ_α's possible execution from s_α. By ($\blacktriangleleft$), all σ_α's complete executions from s_α are finite. Thus, μ is finite. Therefore, $\delta = [t_0] \cdots [t_n]$ for some $n \in \mathbb{N}$.

  We continue to show that $[t_n] \subseteq [\![\varphi]\!]$. Since $\delta = [t_0] \cdots [t_n]$ is a τ_α's complete execution from $[t]$ and it is also a σ_α's possible execution from t, there are two cases: $[t_n] \in \text{CELeaf}(f_\alpha, t')$ for some $t' \in \text{dom}(f_\alpha)$, or δ is a σ_α's complete execution from t. If $[t_n] \in \text{CELeaf}(f_\alpha, t')$ for some $[t'] \in \text{dom}(f_\alpha)$, then there exists $\beta < \alpha$ s.t. $[t] \in \text{CELeaf}(\tau_\beta, t')$ and $[t'] \in \text{dom}(\beta)$. By IH, $[t_n] \subseteq [\![\varphi]\!]$. If δ is a σ_α's complete execution from $[t]$, it follows that μ is a σ_α's complete execution from $[s_\alpha]$. Then by ($\blacktriangleleft$), we have $[t_n] \subseteq [\![\varphi]\!]$.

5. If $[s_i] \notin \text{dom}(\sigma_i)$, it follows by ($\blacktriangleleft$) that $[s_i] \subseteq [\![\varphi]\!]$. Otherwise, there are two cases: $i = 0$ or $i = \alpha > 0$. If $i = 0$, it follows by $[s_0] \in \text{dom}(\sigma_0)$ that $[s_0] \in \text{CEInner}(\sigma_0, s_0)$. Thus, $[s_0] \in \text{dom}(\tau_0)$.

   If $i = \alpha > 0$ and $[s_\alpha] \in \text{dom}(\sigma_\alpha)$, we will show that if $[s_\alpha] \notin \text{dom}(\tau_\alpha)$ then $[s_\alpha] \subseteq [\![\varphi]\!]$. Firstly, we have that $[s_i] \in \text{CEInner}(\sigma_\alpha, s_\alpha)$. Since $[s_\alpha] \notin \text{dom}(\tau_\alpha)$, it follows that $[s_\alpha] \in \text{CELeaf}(f_\alpha, t)$ for some $[t] \in \text{dom}(f_\alpha)$. It follows that there exists $\beta < \alpha$ such that $[s_\alpha] \in \text{CELeaf}(\tau_\beta, t)$ and $t \in \text{dom}(\tau_\beta)$. It follows by (4) that $[s_i] \subseteq [\![\varphi]\!]$. $\blacksquare$

Step II. We define $\tau_\gamma = \bigcup_{i<\gamma} \tau_i$. It follows by (1) and (2) of Claim 3.3.1 that τ_γ is indeed a partial function. Then we prove the following claim.

Claim 3.3.2. *If* $\delta = [t_0] \cdots$ *is a* τ_γ*'s complete execution from* $[t] \in \text{dom}(\tau_\gamma)$ *then* $|\delta| = n$ *for some* $n \in \mathbb{N}$ *and* $[t_n] \subseteq [\![\varphi]\!]$.

Proof of Claim 3.3.2: Since $[t] \in \text{dom}(\tau_\gamma)$, it follows that $[t] \in \text{dom}(\tau_i)$ for some $i < \gamma$. It follows by (5) of Claim 3.3.1 that all τ_i's complete executions from $[t]$ are finite. Thus, there exists $\mu \sqsubseteq \delta$ such that $|\mu| = n$ for some $n \in \mathbb{N}$ and μ is τ_i's complete execution from $[t]$. It follows by (5) of Claim 3.3.1 that $[t_n] \subseteq [\![\varphi]\!]$.

Next, we only need to show $\delta = \mu$. If not, then $\delta = [t_0] \cdots [t_n][t_{n+1}] \cdots$. We then have that there exists $j < \gamma$ such that $\{t_k \mid 0 \leq k \leq n\} \subseteq \text{dom}(\tau_j)$. It cannot be that $j \leq i$. Otherwise, μ is not τ_i's complete execution since $\tau_j \subseteq \tau_i$ by (1) of Claim 3.3.1. Thus, we have $j > i$. Since we also have that $[t_n] \in \text{dom}(\tau_j)$, $[t_n] \in \text{CELeaf}(\tau_i, t)$ and $t \in \text{dom}(\tau_i)$, this is in contradiction with (3) of Claim 3.3.1. Therefore, we have $\delta = \mu$. $\blacksquare$

Step III. We define τ as $\tau = \tau_\gamma \cup (\sigma|_C)$ where $C = \text{CEInner}(\sigma, s) \setminus (\text{dom}(\tau_\gamma) \cup \{[v] \in \text{CELeaf}(\tau', t) \mid [t] \in \text{dom}(\tau_\gamma)\})$ and σ is the strategy mentioned at ($*$). Since both τ_γ and $\sigma|_C$ are partial functions, τ is also a partial function. We then prove the following claim.

Claim 3.3.3. *If* $\delta = [t_0] \cdots$ *is a* τ*'s complete execution from* $[t] \in \text{dom}(\tau)$ *then* $|\delta| = n$ *for some* $n \in \mathbb{N}$ *and* $[t_n] \subseteq [\![\varphi]\!]$.

Proof of Claim 3.3.3: Since $\text{dom}(\tau) = \text{dom}(\tau_\gamma) \cup C$, there are two cases: $[t] \in \text{dom}(\tau_\gamma)$ or $[t] \in C$.

If $[t] \in \text{dom}(\tau_\gamma)$, it follows that $\text{CELeaf}(\tau_\gamma, t) \cap C = \emptyset$. Moreover, we have $\text{CELeaf}(\tau_\gamma, t) \cap \text{dom}(\tau_\gamma) = \emptyset$. Thus, we have $\text{CELeaf}(\tau_\gamma, t) \cap \text{dom}(\tau) = \emptyset$. It follows by Proposition 3.2 that δ is a τ_γ's complete execution from from $[t]$. It follows by Claim 3.3.2 $|\delta| = n$ for some $n \in \mathbb{N}$ and $[t_n] \subseteq [\![\varphi]\!]$.

If $[t] \in C$, there are two cases: there exists $k < |\delta|$ such that $[t_k] \in \text{dom}(\tau_\gamma)$, or there does not exists such k. (Please note that $|\delta| > 1$ due to the fact that $\delta = [t_0] \cdots$ is τ's complete execution from $[t] \in \text{dom}(\tau)$).

- $[t_k] \in \text{dom}(\tau_\gamma)$ for some $k < |\delta|$: It follows that $\mu = [t_k] \cdots$ is a τ's complete execution from $[t_k]$. Since $\text{dom}(\tau) \cap \text{CELeaf}(\tau_\gamma, t_k) = \emptyset$, it follows by Proposition 3.2 that μ is a τ_γ's complete execution from $[t_k]$. It follows by Claim 3.3.2 that $\mu = [t_k] \cdots [t_{k+n}]$ for some $n \in \mathbb{N}$ and $[t_{k+n}] \subseteq [\![\varphi]\!]$. Therefore, $|\delta| = k + n$.

- If there does not exist $k < |\delta|$ s.t. $[t_k] \in \text{dom}(\tau_\gamma)$, then $\delta = [t_0] \cdots$ is a σ's possible execution from $[t]$. Since $[t] \in C \subseteq \text{CEInner}(\sigma, s)$, then there is a σ's possible execution $[s_0] \cdots [s_m]$ s.t. $m \in \mathbb{N}$, $[s_0] = [s]$ and $[s_m] = [t]$. Let $\mu = [s_0] \cdots [s_{m-1}] \circ \delta$. (If $m = 0$ then $\mu = \delta$). It follows that μ is σ's possible execution from s. By ($*$), all σ's complete executions from s are finite. Thus, μ is finite. Therefore, $\delta = [t_0] \cdots [t_n]$ for some $n \in \mathbb{N}$.

  We continue to show that $[t_n] \subseteq [\![\varphi]\!]$. Since $\delta = [t_0] \cdots [t_n]$ is a τ's complete execution from t and it is also a σ's possible execution from t, there are two cases: $[t_n] \in \text{CELeaf}(\tau_\gamma, t')$ for some $t' \in \text{dom}(\tau_\gamma)$, or δ is a σ's complete execution from t. If $[t_n] \in \text{CELeaf}(\tau_\gamma, t')$ for some $[t'] \in \text{dom}(\tau_\gamma)$, it follows by Claim 3.3.2 that $[t_n] \subseteq [\![\varphi]\!]$. If δ is a σ's complete execution from $[t]$, it follows μ is a σ's complete execution from $[s]$. It follows that $[t_n] = S_i$ for some $0 \leq i < \gamma$. Since $\delta = [t_0] \cdots [t_n]$ is τ's complete execution from $[t] \in \text{dom}(\tau_\gamma)$, it follows $[t_n] \notin \text{dom}(\tau_\gamma)$. We then have $[t_n] \notin \text{dom}(\tau_i)$, namely $S_i \notin \tau_i$. It follows by (5) of Claim 3.3.1 that $S_i \subseteq [\![\varphi]\!]$, namely $[t_n] \subseteq [\![\varphi]\!]$. $\blacksquare$

Next, we continue to show that $\mathcal{M}, s \vDash \mathcal{K}h\varphi$ with the assumption that $[s] \in \text{dom}(\sigma)$. Since $[s] \in \text{dom}(\sigma)$, we have $[s] \in \text{CEInner}(\sigma, s)$. There are two cases: $[s] \in \text{dom}(\tau)$ or not. If $[s] \in \text{dom}(\tau)$, it follows by Claim 3.3.3 that $\mathcal{M}, s \vDash \mathcal{K}h\varphi$. If $[s] \notin \text{dom}(\tau)$, due to $[s] \in \text{CEInner}(\sigma, s)$, it follows that $[s] \in \text{CELeaf}(\tau_\gamma, t)$ for some $[t] \in \text{dom}(\tau_\gamma)$. It follows by Claim 3.3.2 that $[s] \subseteq [\![\varphi]\!]$. It follows that $\mathcal{M}, s \vDash \mathcal{K}\varphi$. It is obvious that $\mathcal{M}, s \vDash \mathcal{K}h\varphi$. $\square$

The validity of the rest of the axioms can be checked (cf. the explanation right after the system $\mathbb{SKH}$), then we have:

Theorem 3.4 (Soundness). *If $\vdash \varphi$ then $\vDash \varphi$.*

4 Completeness and Decidability

Let Φ be a subformula-closed set of formulas. It is obvious that Φ is countable since the whole language itself is countable. Given a set of formulas Δ, let: $\Delta|_{\mathcal{K}} = \{\mathcal{K}\varphi \mid \mathcal{K}\varphi \in \Delta\}$, $\Delta|_{\neg\mathcal{K}} = \{\neg\mathcal{K}\varphi \mid \neg\mathcal{K}\varphi \in \Delta\}$, $\Delta|_{\mathcal{K}h} = \{\mathcal{K}h\varphi \mid \mathcal{K}h\varphi \in \Delta\}$, $\Delta|_{\neg\mathcal{K}h} = \{\neg\mathcal{K}h\varphi \mid \neg\mathcal{K}h\varphi \in \Delta\}$. Below we define the closure of Φ, and use it to build a canonical model w.r.t. Φ. We will show that when Φ is finite then we can build a finite model.

Definition 4.1. *The closure $cl(\Phi)$ is $\Phi \cup \{\mathcal{K}\varphi \mid \varphi \in \Phi\}$.*

Definition 4.2 (Atom). *We enumerate the formulas in $cl(\Phi)$ by $\{\psi_i \mid i \in \mathbb{N}\}$. The formula set $\Delta = \{Y_i \mid i \in \mathbb{N}\}$ is an atom of $cl(\Phi)$ if*

- $Y_i = \psi_i$ or $Y_i = \neg\psi_i$ *for each $\psi_i \in cl(\Phi)$;*

- Δ *is consistent.*

Note that if Φ is the whole language then an atom is simply a maximal consistent set. By a standard inductive construction, we can obtain the Lindenbaum-like result in our setting (which is useful to show the existence lemma for $\mathcal{K}$):

Proposition 4.3. *Let Δ be an atom of $cl(\Phi)$, $\Gamma \subseteq \Delta$ and $\varphi \in cl(\Phi)$. If $\Gamma \cup \{\pm\varphi\}$ is consistent then there is an atom Δ' of $cl(\Phi)$ such that $(\Gamma \cup \{\pm\varphi\}) \subseteq \Delta'$, where $\pm\varphi = \varphi$ or $\pm\varphi = \neg\varphi$.*

Definition 4.4. *Given a subformula-closed Φ, the canonical model $\mathcal{M}^{\Phi} = \langle W, ACT, \sim, \{\xrightarrow{x} \mid x \in ACT\}, V\rangle$ is defined as:*

- $W = \{\Delta \mid \Delta$ *is an atom of $cl(\Phi)\}$;*

- $ACT = \{\varphi \mid \mathcal{K}h\varphi \in \Phi\}$;

- $\Delta \sim \Delta'$ *iff $\Delta|_{\mathcal{K}} = \Delta'|_{\mathcal{K}}$;*

- *for each $\varphi \in ACT$, $\Delta \xrightarrow{\varphi} \Delta'$ iff $\mathcal{K}h\varphi, \neg\mathcal{K}\varphi \in \Delta$ and $\mathcal{K}\varphi \in \Delta'$;*

- *for each $p \in \Phi$, $p \in V(\Delta)$ iff $p \in \Delta$.*

Note that we use the φ in $\mathcal{K}h\varphi$ formulas as the action labels, and we introduce an action transition if it is necessary, i.e., $\mathcal{K}h\varphi$ but $\neg\mathcal{K}\varphi$ (i.e., empty strategy does not work). Requiring $\mathcal{K}\varphi \in \Delta'$ is to reflect the first condition in the semantics of $\mathcal{K}h$. Using NECK, DISTK and Proposition 4.3, it is routine to show the existence lemma for $\mathcal{K}$:

Proposition 4.5. *Let Δ be a state in $\mathcal{M}^{\Phi}$, and $\mathcal{K}\varphi \in cl(\Phi)$. If $\mathcal{K}\varphi \notin \Delta$ then there exists $\Delta' \in [\Delta]$ such that $\neg\varphi \in \Delta'$.*

Proof. Let $\Gamma = \Delta|_{\mathcal{K}} \cup \Delta|_{\neg\mathcal{K}} \cup \{\neg\varphi\}$. Γ is consistent. If not, there are $\mathcal{K}\varphi_i, \cdots, \mathcal{K}\varphi_n$ and $\neg\mathcal{K}\psi_1, \cdots, \neg\mathcal{K}\psi_m$ in Δ:

$$\vdash \mathcal{K}\varphi_1 \wedge \cdots \wedge \mathcal{K}\varphi_n \wedge \neg\mathcal{K}\psi_1 \wedge \cdots \wedge \neg\mathcal{K}\psi_m \to \varphi.$$

Following by NECK and DISTK, we have

$$\vdash \mathcal{K}(\mathcal{K}\varphi_i \wedge \cdots \wedge \mathcal{K}\varphi_n \wedge \neg\mathcal{K}\psi_1 \wedge \cdots \wedge \neg\mathcal{K}\psi_m) \to \mathcal{K}\varphi.$$

Since the epistemic operator $\mathcal{K}$ is distributive over $\wedge$ and $\vdash \mathcal{K}\mathcal{K}\varphi_i \leftrightarrow \mathcal{K}\varphi_i$ for all $1 \le i \le n$ and $\vdash \mathcal{K}\neg\mathcal{K}\psi_i \leftrightarrow \neg\mathcal{K}\psi_i$ for all $1 \le i \le m$, we have

$$\vdash \mathcal{K}\varphi_i \wedge \cdots \wedge \mathcal{K}\varphi_n \wedge \neg\mathcal{K}\psi_1 \wedge \cdots \wedge \neg\mathcal{K}\psi_m \to \mathcal{K}\varphi.$$

Since $\mathcal{K}\varphi_i, \cdots, \mathcal{K}\varphi_n$ and $\neg\mathcal{K}\psi_1, \cdots, \neg\mathcal{K}\psi_m$ are all in Δ and $\mathcal{K}\varphi \in cl(\Phi)$, it follows that $\mathcal{K}\varphi \in \Delta$. It is contradictory with the assumption that $\mathcal{K}\varphi \notin \Delta$. Therefore, Γ is consistent. It follows by Proposition 4.3 that there exists an atom Δ' of $cl(\Phi)$ such that $\Gamma \subseteq \Delta'$. Since $(\Delta|_{\mathcal{K}} \cup \Delta|_{\neg\mathcal{K}}) \subseteq \Delta'$, we have $\Delta' \sim \Delta$, that is, $\Delta' \in [\Delta]$. $\square$

Proposition 4.6. *Let Δ and Δ' be two states in $\mathcal{M}^{\Phi}$ such that $\Delta \sim \Delta'$. We have $\Delta|_{\mathcal{K}h} = \Delta'|_{\mathcal{K}h}$.*

Proof. For each $\mathcal{K}h\varphi \in \Delta$, by Definition 4.1, $\mathcal{K}h\varphi \in \Phi$. Then $\mathcal{K}\mathcal{K}h\varphi \in cl(\Phi)$. For each $\mathcal{K}h\varphi \in \Delta$, by Axiom AxKhtoKKh, we have $\mathcal{K}\mathcal{K}h\varphi \in \Delta$. Since $\Delta \sim \Delta'$, then $\mathcal{K}\mathcal{K}h\varphi \in \Delta'$, and by Axiom T, $\mathcal{K}h\varphi \in \Delta'$. Then we showed that $\mathcal{K}h\varphi \in \Delta$ implies $\mathcal{K}h\varphi \in \Delta'$. Similarly we can prove $\mathcal{K}h\varphi \in \Delta'$ implies $\mathcal{K}h\varphi \in \Delta$. Hence, $\Delta|_{\mathcal{K}h} = \Delta'|_{\mathcal{K}h}$. $\square$

The following is a crucial observation for the later proofs.

Proposition 4.7. *Let Δ be a state in $\mathcal{M}^{\Phi}$ and $\psi \in ACT$ be executable at $[\Delta]$. If $\mathcal{K}h\varphi \in \Delta'$ for all Δ' with $[\Delta] \xrightarrow{\psi} [\Delta']$ then $\mathcal{K}h\varphi \in \Delta$.*

Proof. First, we show that $\mathcal{K}\psi$ is not consistent with $\neg\mathcal{K}h\varphi$. It is obvious that $\mathcal{K}h\varphi \in cl(\Phi)$. Since ψ is executable at $[\Delta]$, there are atoms Γ_1 and Γ_2 s.t. $\Gamma_1 \xrightarrow{\psi} \Gamma_2$. Then $\mathcal{K}\psi \in \Gamma_2$. Assuming that $\mathcal{K}\psi$ is consistent with $\neg\mathcal{K}h\varphi$, by Proposition 4.3 there exists an atom Γ of $cl(\Phi)$ s.t. $\{\mathcal{K}\psi, \neg\mathcal{K}h\varphi\} \subseteq \Gamma$. Since $\psi \in ACT$ is executable at $[\Delta]$, then by definition of $\xrightarrow{\psi}, \sim$ and Proposition 4.6, $\mathcal{K}h\psi, \neg\mathcal{K}\psi \in \Delta$. It follows that $\Delta \xrightarrow{\psi} \Gamma$, then $[\Delta] \xrightarrow{\psi} [\Gamma]$. This is in contradiction with the assumption that $\mathcal{K}h\varphi \in \Delta'$ for all Δ' with $[\Delta] \xrightarrow{\psi} [\Delta']$. Then $\mathcal{K}\psi$ is not consistent with $\neg\mathcal{K}h\varphi$. Hence, $\vdash \mathcal{K}\psi \to \mathcal{K}h\varphi$.

Since $\vdash \mathcal{K}\psi \to \mathcal{K}h\varphi$, it follows by Rule MONOKh and Axiom AxKhtoKhK that $\vdash \mathcal{K}h\psi \to \mathcal{K}h\mathcal{K}h\varphi$. Moreover, it follows by Axiom AxKhKh that $\vdash \mathcal{K}h\psi \to \mathcal{K}h\varphi$. Since ψ is executable at $[\Delta]$, it follows by the definition of $\xrightarrow{\psi}$ and Proposition 4.6 that $\mathcal{K}h\psi \in \Delta$. Therefore, we have $\mathcal{K}h\varphi \in \Delta$. $\square$

Lemma 4.8. *For each $\varphi \in cl(\Phi)$, $\mathcal{M}^{\Phi}, \Delta \vDash \varphi$ iff $\varphi \in \Delta$.*

Proof. We prove it by induction on φ. We only focus on the case of $\mathcal{K}h\varphi \in cl(\Phi)$; the other cases are straightforward, e.g., $\mathcal{K}\varphi$ case can be proved based on Proposition 4.5. Note that if $\mathcal{K}h\varphi \in cl(\Phi)$ then $\mathcal{K}h\varphi \in \Phi$, therefore $\varphi \in \Phi$ since Φ is subformula-closed. Thus by Definition 4.1, $\mathcal{K}\varphi \in cl(\Phi)$.

Right to Left: If $\mathcal{K}h\varphi \in \Delta$, we will show $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}h\varphi$. Firstly, there are two cases: $\mathcal{K}\varphi \in \Delta$ or $\mathcal{K}\varphi \notin \Delta$. If $\mathcal{K}\varphi \in \Delta$, then $\mathcal{K}\varphi, \varphi \in \Delta'$ for all $\Delta' \in [\Delta]$. Since $\varphi \in \Phi$, it follows by IH that $\mathcal{M}^{\Phi}, \Delta' \vDash \varphi$ for all $\Delta' \in [\Delta]$. Therefore, $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}\varphi$. It follows by Axiom AxKtoKh and the soundness of $\mathbb{SKH}$ that $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}h\varphi$. If $\neg\mathcal{K}\varphi \in \Delta$, we first show that $\mathcal{K}\varphi$ is consistent. If not, namely $\vdash \mathcal{K}\varphi \to \bot$, it follows by Rule MONOKh that $\vdash \mathcal{K}h\mathcal{K}\varphi \to \mathcal{K}h\bot$. It follows by Axiom AxKhbot that $\vdash \mathcal{K}h\mathcal{K}\varphi \to \bot$. Since $\mathcal{K}h\varphi \in \Delta$, it follows by Axiom AxKhtoKhK that $\Delta \vdash \bot$, which is in contradiction with the fact that Δ is consistent. Therefore, $\mathcal{K}\varphi$ is consistent.

By Proposition 4.3 there exists an atom Δ' s.t. $\mathcal{K}\varphi \in \Delta'$. Note that $\varphi \in$ ACT. Thus, we have $\Delta \xrightarrow{\varphi} \Delta'$, then $[\Delta] \xrightarrow{\varphi} [\Delta']$. Let $[\Delta'']$ be an equivalence class s.t. $[\Delta] \xrightarrow{\varphi} [\Delta'']$, which indicates $\Gamma \xrightarrow{\varphi} \Gamma''$ for some $\Gamma \in [\Delta]$ and $\Gamma'' \in [\Delta'']$. By definition of $\xrightarrow{\varphi}$ and $\sim$ we get $\mathcal{K}\varphi \in \Theta$ for all $\Theta \in [\Delta'']$. By IH, $\mathcal{M}^{\Phi}, \Theta \vDash \varphi$ for all $\Theta \in [\Delta'']$, namely $[\Delta''] \subseteq \llbracket \varphi \rrbracket$. Moreover, $\xrightarrow{\varphi}$ is not a loop on $[\Delta]$ because $\neg \mathcal{K}\varphi \in \Delta$. Thus, the partial function $\sigma = \{[\Delta] \mapsto \varphi\}$ is a strategy s.t. all its complete executions starting from $[\Delta]$ are finite and $[\Delta''] \subseteq \llbracket \varphi \rrbracket$ for each $[\Delta''] \in \texttt{CELeaf}(\sigma, \Delta)$. Then, $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}h\varphi$.

Left to Right: Suppose $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}h\varphi$, we will show $\mathcal{K}h\varphi \in \Delta$. By the semantics, there exists a strategy σ s.t. all σ's complete executions starting from $[\Delta]$ are finite and $[\Gamma] \subseteq \llbracket \varphi \rrbracket$ for all $[\Gamma] \in \texttt{CELeaf}(\sigma, \Delta)$. By IH, $\varphi \in \Gamma'$ for all $\Gamma' \in [\Gamma]$ and $[\Gamma] \in \texttt{CELeaf}(\sigma, \Delta)$. By Proposition 4.5, we get $\mathcal{K}\varphi \in \Gamma$ for all $[\Gamma] \in \texttt{CELeaf}(\sigma, \Delta)$. By Axiom AxKtoKh and Proposition 4.6, $\mathcal{K}h\varphi \in \Gamma$ for all $[\Gamma] \in \texttt{CELeaf}(\sigma, \Delta)$.

If $[\Delta] \notin \texttt{dom}(\sigma)$, it is obvious that $\mathcal{K}h\varphi \in \Delta$ because $[\Delta] \in \texttt{CELeaf}(\sigma, \Delta)$. Next, we consider the case of $[\Delta] \in \texttt{dom}(\sigma)$, then $[\Delta] \in \texttt{CEInner}(\sigma, \Delta)$. In order to show $\mathcal{K}h\varphi \in \Delta$, we will show a stronger result that $\mathcal{K}h\varphi \in \Delta'$ for all $[\Delta'] \in \texttt{CEInner}(\sigma, \Delta)$. Firstly, we show the following claim:

Claim 4.8.1. *If there exists* $[\Delta'] \in \texttt{CEInner}(\sigma, \Delta)$ *such that* $\neg \mathcal{K}h\varphi \in \Delta'$ *then there exists an infinite execution of* σ *starting from* $[\Delta]$.

Proof of Claim 4.8.1: Let X be the set $\{[\Theta] \in \texttt{CEInner}(\sigma, \Delta) \mid \neg \mathcal{K}h\varphi \in \Theta\}$. It follows that $[\Delta'] \in X$ and $X \subseteq \texttt{dom}(\sigma)$. We define a binary relation R on X as $R = \{([\Theta], [\Theta']) \mid [\Theta] \xrightarrow{\sigma([\Theta])} [\Theta']\}$.

For each $[\Theta] \in X$, we have that $\sigma([\Theta])$ is executable at $[\Theta]$. Since $\neg \mathcal{K}h\varphi \in \Theta$, by Proposition 4.7 there exists an atom Θ' s.t. $[\Theta] \xrightarrow{\sigma([\Theta])} [\Theta']$ and $\neg \mathcal{K}h\varphi \in \Theta'$. Since $\mathcal{K}h\varphi \in \Gamma$ for all $[\Gamma] \in \texttt{CELeaf}(\sigma, \Delta)$ and $[\Theta] \in \texttt{CEInner}(\sigma, \Delta)$, we have $[\Theta'] \in \texttt{CEInner}(\sigma, \Delta)$. Then $[\Theta'] \in X$. Therefore, R is an entire binary relation on X, namely for each $[\Theta] \in X$ there is $[\Theta'] \in X$ such that $([\Theta], [\Theta']) \in R$. Then by Axiom of Dependent Choice there exists an infinite sequence $[\Theta_0][\Theta_1] \cdots$ s.t. $([\Theta_n], [\Theta_{n+1}]) \in R$ for all $n \in \mathbb{N}$.

From the definition of R, $[\Theta_0][\Theta_1] \cdots$ is a complete execution of σ starting from $[\Theta_0]$. Since $[\Theta_0] \in \texttt{CEInner}(\sigma, \Delta)$ and all complete execution of σ from $[\Delta]$ are finite, there is a possible execution $[\Delta_0] \cdots [\Delta_j]$ for some $j \in \mathbb{N}$ s.t. $[\Delta_0] = [\Delta]$ and $[\Delta_j] = [\Theta_0]$. Therefore, $[\Delta_0] \cdots [\Delta_j][\Theta_1] \cdots$ is an infinite complete execution of σ from $[\Delta]$. ∎

Therefore, we have $\mathcal{K}h\varphi \in \Delta'$ for all $[\Delta'] \in \texttt{CEInner}(\sigma, s)$. Otherwise, by Claim 4.8.1 there is an infinite complete execution given σ from $[\Delta]$. This is contradictory with all σ's complete execution from $[\Delta]$ are finite, then $\mathcal{K}h\varphi \in \Delta'$ for all $[\Delta'] \in \texttt{CEInner}(\sigma, s)$. Since $[\Delta] \in \texttt{dom}(\sigma)$, we get $[\Delta] \in \texttt{CEInner}(\sigma, \Delta)$. Then $\mathcal{K}h\varphi \in \Delta$. □

Now let Φ be the set of all formulas, then each maximal consistent set Δ is actually an atom which satisfies all its formulas in $\mathcal{M}^{\Phi}$, according to the above truth lemma. Completeness then follows immediately.

Theorem 4.9. $\mathbb{SKH}$ *is strongly complete.*

Note that if Φ is the set of all subformulas of a given formula φ, then $cl(\Phi)$ is still finite. Due to the soundness of $\mathbb{SKH}$ and Proposition 4.3, a satisfiable formula φ must be consistent thus appearing in some atom, and thus φ is satisfiable in $\mathcal{M}^{\Phi}$. It is not hard to see that $|\mathcal{M}^{\Phi}| \leq 2^{2|\varphi|}$ where $2|\varphi|$ is the bound on the size of $cl(\Phi)$. This gives us a small model property of our logic, then decidability follows.

Theorem 4.10. $\mathbb{SKH}$ *is decidable.*

5 Conclusion

In this paper, we propose an epistemic logic of both (goal-directed) *knowing how* and *knowing that*, and capture their interactions. A sound and complete proof system $\mathbb{SKH}$ is obtained, and we prove the decidability of the logic. We hope the axioms are illuminating towards a better understanding of *knowing how*.

Note that we do not impose any special properties between the interaction of $\xrightarrow{a}$ and $\sim$ in the models so far. In the future, it would be interesting to see whether assuming properties of *perfect recall* and/or *no learning* (cf. e.g., [Fagin *et al.*, 1995; Wang and Li, 2012]) changes the logic or not.

Our notion of knowing how is relatively strong, particularly evidenced by the axiom AxKhtoKhK : $\mathcal{K}h\varphi \to \mathcal{K}h\mathcal{K}\varphi$, which is due to the first condition of our semantics for $\mathcal{K}h$, inspired by planning with uncertainty. We believe it is reasonable for the scenarios where the agent has perfect recall (or, say, never forgets), which is usually assumed implicitly in the discussions on planning (cf. [Li *et al.*, 2017]). However, for a forgetful agent it may not be intuitive anymore, e.g., I know how to get drunk when sober but I may not know how to get to the state that I know I am drunk, assuming drunk people do not know they are drunk. The axiom AxKhKh is also interesting in distinguishing different types of knowing how. We have been focusing on the goal-directed knowing how [Gochet, 2013], but for other types of knowing how such as knowing how to swim, the axiom may not be reasonable anymore, e.g., I know how to let myself to know how to swim (by registering an excellent swimming course) does not mean that I know how to swim right now. We leave the discussion of other types of knowing how in the future. Moreover, by introducing suitable notions of bisimulations we will be able to compare their expressive power. Another obvious next step is to consider knowing how in multi-agent settings, which brings us closer to coalition logics (cf. e.g., [Ågotnes and Alechina, 2012; Naumov and Tao, 2017].

References

[Ågotnes and Alechina, 2012] Thomas Ågotnes and Natasha Alechina. Epistemic coalition logic: completeness and complexity. In *Proceedings of AAMAS 2012*, pages 1099–1106, 2012.

[Ågotnes *et al.*, 2015] Thomas Ågotnes, Valentin Goranko, Wojciech Jamroga, and Michael Wooldridge. Knowledge and ability. In Hans van Ditmarsch, Joseph Halpern, Wiebe van der Hoek, and Barteld Kooi, editors, *Handbook of Epistemic Logic*, chapter 11, pages 543–589. College Publications, 2015.

[Belardinelli, 2014] Francesco Belardinelli. Reasoning about knowledge and strategies: Epistemic strategy logic. In *Proceedings of SR 2014*, pages 27–33, 2014.

[Broersen and Herzig, 2015] Jan Broersen and Andreas Herzig. Using STIT theory to talk about strategies. In Johan van Benthem, Sujata Ghosh, and Rineke Verbrugge, editors, *Models of Strategic Reasoning: Logics, Games, and Communities*, pages 137–173. Springer, 2015.

[Fagin *et al.*, 1995] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning about knowledge*. MIT Press, 1995.

[Fantl, 2008] Jeremy Fantl. Knowing-how and knowing-that. *Philosophy Compass*, 3(3):451–470, 2008.

[Geffner and Bonet, 2013] Hector Geffner and Blai Bonet. A concise introduction to models and methods for automated planning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 8(1):1–141, 2013.

[Gochet, 2013] Paul Gochet. An open problem in the logic of knowing how. In J. Hintikka, editor, *Open Problems in Epistemology*. The Philosophical Society of Finland, 2013.

[Herzig *et al.*, 2013] Andreas Herzig, Emiliano Lorini, and Dirk Walther. Reasoning about actions meets strategic logics. In *Proceedings of LORI 2013*, pages 162–175, 2013.

[Herzig, 2015] Andreas Herzig. Logics of knowledge and action: critical analysis and challenges. *Autonomous Agents and Multi-Agent Systems*, 29(5):719–753, 2015.

[Hintikka, 1962] J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca N.Y., 1962.

[Jamroga and Ågotnes, 2007] Wojciech Jamroga and Thomas Ågotnes. Constructive knowledge: what agents can achieve under imperfect information. *Journal of Applied Non-Classical Logics*, 17(4):423–475, 2007.

[Lespérance *et al.*, 2000] Yves Lespérance, Hector J. Levesque, Fangzhen Lin, and Richard B. Scherl. Ability and knowing how in the situation calculus. *Studia Logica*, 66(1):165–186, 2000.

[Li and Wang, 2017] Yanjun Li and Yanjing Wang. Achieving while maintaining: A logic of knowing how with intermediate constraints. In *Proceedings of ICLA'17*, pages 154–167, 2017.

[Li *et al.*, 2017] Yanjun Li, Quan Yu, and Yanjing Wang. More for free: a dynamic epistemic framework for conformant planning over transition systems. *Journal of Logic and Computation*, 2017. forthcoming.

[Li, 2017] Yanjun Li. Stopping means achieving: A weaker logic of knowing how. *Studies in Logic*, 9(4):34–54, 2017.

[McCarthy and Hayes, 1969] John McCarthy and Patrick J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence*, pages 463–502. Edinburgh University Press, 1969.

[McCarthy, 1979] John McCarthy. First-Order theories of individual concepts and propositions. *Machine Intelligence*, 9.:129–147, 1979.

[Moore, 1985] Robert C Moore. A formal theory of knowledge and action. In Jerry R. Hobbs and Robert C. Moore, editors, *Formal Theories of the Commonsense World*. Ablex Publishing Corporation, 1985.

[Naumov and Tao, 2017] Pavel Naumov and Jia Tao. Coalition power in epistemic transition systems. In *Proceedings of AAMAS 2017*, pages 723–731, 2017.

[Smith and Weld, 1998] David E. Smith and Daniel S. Weld. Conformant graphplan. In *AAAI 98*, pages 889–896, 1998.

[Stanley and Williamson, 2001] Jason Stanley and Timothy Williamson. Knowing how. *The Journal of Philosophy*, pages 411–444, 2001.

[van der Hoek and Wooldridge, 2003] Wiebe van der Hoek and Machael Wooldridge. Cooperation, knowledge, and time: Alternating-time temporal epistemic logic and its applications. *Studia Logica*, (1):125–157, 2003.

[van der Hoek *et al.*, 2000] Wiebe van der Hoek, Bernd van Linder, and John-Jules Ch. Meyer. On agents that have the ability to choose. *Studia Logica*, 66(1):79–119, 2000.

[Wang and Li, 2012] Yanjing Wang and Yanjun Li. Not all those who wander are lost: Dynamic epistemic reasoning in navigation. In *Proceedings of AiML 2012*, pages 559–580, 2012.

[Wang, 2015] Yanjing Wang. A logic of knowing how. In *Proceedings of LORI 2015*, pages 392–405, 2015.

[Wang, 2016a] Yanjing Wang. Beyond knowing that: a new generation of epistemic logics. In Hans van Ditmarsch and Gabriel Sandu, editors, *Jaakko Hintikka on knowledge and game theoretical semantics*. Springer, 2016. forthcoming.

[Wang, 2016b] Yanjing Wang. A logic of goal-directed knowing how. *Synthese*, pages 1–21, 2016. in press.

[Yu *et al.*, 2016] Quan Yu, Yanjun Li, and Yanjing Wang. A dynamic epistemic framework for conformant planning. In *Proceedings of TARK'15*, pages 298–318. EPTCS, 2016.