Proceedings of the 22nd International Conference on Principles of Knowledge Representation and Reasoning
Main Track

How Lucky Are You to Know Your Way? A Probabilistic Approach to Knowing How Logics

Pablo F. Castro[1,3] , **Pedro R. D'Argenio**[2,3] , **Raul Fervari**[2,3,4]
[1]Universidad Nacional de Río Cuarto, FCEFQyN, Departamento de Computación, Argentina
[2]Universidad Nacional de Córdoba, FAMAF, Argentina
[3]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina
[4]Université Paris-Saclay, CNRS, ENS Paris-Saclay, Laboratoire Méthodes Formelles, France

Abstract

We introduce a probabilistic version of knowing-how modal logics. More precisely, our logics extend extant approaches to model the ability of an agent to achieve a given goal with a certain probability. On the semantic side, we enrich the models of the logic with probability distributions over the agent's actions. Then, we investigate different languages to describe such structures. First, we consider a probabilistic version of the linear plan-based logic of knowing how, and discuss its properties. Then, we consider indistinguishability classes, and obtain two logics, one that has 'non-adaptative' plans, and another with 'adaptative' plans. In all cases we investigate the computational complexity of their model-checking problem, obtaining undecidability results for the first and the second logic, while for the last one the problem is decidable in polynomial time. We also explore the semantics of the new logics under non-probabilistic models to compare them to the original non-probabilistic ones.

1 Introduction

Modern approaches in Epistemic Logic (Hintikka 1962) have shifted focus from a single notion of knowledge (usually, the notion of *knowing that*) to a diverse palette of notions, each of them tailored to specific purposes. In this regard, the notion of *knowing how* has received significant attention, since it captures scenarios related to intelligent agents and its strategic behaviour. Logically, knowing how is typically defined as the ability of an agent to achieve a certain goal. Knowing-how logics have direct applications to *planning* problems (Russell and Norvig 2020), where plans have to be constructed such that a collection of agents can achieve a given goal. Examples of planning applications can be found in e.g. self driving cars, robotics, conversational agents, cybersecurity, risk management, etc.

Usually, the semantics for knowing-how logics can be thought as a combination of operators that describe abilities alongside standard epistemic operators for knowing that. This is the approach introduced in, e.g. (McCarthy and Hayes 1969; Moore 1985; Lespérance et al. 2000; van der Hoek, van Linder, and Meyer 2000; Herzig and Troquard 2006). As a result, knowing how reflects that *the agent knows that there is a course of action leading to achieve the intended goal*. However, as pointed out in e.g. (Jamroga and Ågotnes 2007; Herzig 2015), this reading is not entirely accurate. Instead, knowing how could be read as *there is a course of action, that the agent knows how to apply, to bring about the goal*. Thus, a novel perspective emerged in (Wang 2015; Wang 2018a; Wang 2018b), where a new modality is defined with the aim of capturing the proper reading of knowing how.

More specifically, the new modality under consideration is a binary modality $\mathsf{Kh}(\psi, \varphi)$ interpreted over Labeled Transition Systems (LTS), where an LTS models the actions that are available to the agents as well as the effects of these actions. Thus, the formula $\mathsf{Kh}(\psi, \varphi)$ holds if there exists a sequence of actions π (i.e., a *plan*) such that, in every situation where ψ holds, π can be executed, it never aborts its execution, and it always leads to situations in which φ holds. This new view of knowing how raised a new family of logics refining the original one, witnessed by the extensive related literature (see e.g. (Li and Wang 2017; Li 2017; Fervari et al. 2017; Naumov and Tao 2019; Naumov and Tao 2018)). Interestingly, in (Areces et al. 2021; Areces et al. 2025) the original approach is enriched by a notion of 'epistemic indistinguishability' between plans, arguably closer to standard semantics of epistemic logics. This indistinguishability relation indicates that all the related plans are considered or perceived "equally good" from the agent's perspective (even if they are not), thus a plan π is suitable for achieving a goal if this is also the case for all the plans that are indistinguishable from π. The new semantics arguably offers a more adequate view of knowing how from an epistemic perspective, compared to the original approach.

The above-mentioned works investigate various logical properties, including axiomatizations, expressivity, and the complexity of the respective logics. In particular, the model-checking problem results of interest, since it is ubiquitous in software verification but also an important tool for controller synthesis and planning. Moreover, as argued in (Demri and Fervari 2023), the model-checking problem better reflects the real power of the logics. This is because these logics often have a simple syntax combined with a rich semantics. In model-checking, plans are part of the input, so the complexity needs to be tamed (unlike in the satisfiability problem, where other tricks can be used like guessing a proper plan). Therein, the authors also discuss how to incorporate other constraints into the plans, more precisely, the different semantics of knowing how modalities are enriched with regularity constraints (i.e., where plans are given by some regular

formalism) and numerical constraints (i.e., where actions in knowing how are restricted by some budget).

The work in (Demri and Fervari 2023) paves the way for studying additional constraints, particularly knowing how to achieve a goal with a certain probability. This is of interest in scenarios where plans might lead to unexpected results due to a faulty behavior of actions, or because their executions lead to random outcomes. Just as (constrained) planning is connected to (constrained) knowing-how, the ability to handle probabilities in the context of knowing how serves as the logical counterpart to probabilistic planning (see, e.g., (Madani, Hanks, and Condon 1999)) and related concepts. Moreover, there is a realm of logics featuring probabilistic notions of strategic reasoning. For instance, (Baier et al. 2018) discusses the idea of model-checking with probabilities, while those specifically related to ATL are explored in (Bianco and de Alfaro 1995; Chen and Lu 2007; Bulling and Jamroga 2009). Also, probabilistic strategy logics are investigated in (Aminof et al. 2019), while (Berthon et al. 2024) considers stochastic natural strategic abilities.

Here, we present extensions of knowing how logics with probabilities. The new modality, written $\mathsf{Kh}_q(\psi, \varphi)$, will be read as *the agent knows how to achieve φ given ψ, with a probability of at least q*. This idea results helpful in modeling case studies where the result of actions have a random component. A simple example of this situation is given in (Kushmerick, Hanks, and Weld 1995): consider a robot that has to grasp an object, the result of the robot's actions stochastically depends on the state of the world. For instance, if the gripper is wet, there is a probability of 0.9 that the object falls when the robot tries to pick it up. Thus, the robot may try to dry the gripper before picking up elements. These kinds of scenarios can be modeled with Probabilistic LTS (PLTS), i.e., transition systems where now transitions relate states with probability distributions, which in turn capture the stochastic behavior of actions.

We start in Sec. 4.1 by naturally extending the logic over linear plans from (Wang 2015; Wang 2018a; Wang 2018b). For this logic we show that, under non-probabilistic models (i.e., LTSs), it agrees with the non-probabilistic case. We prove that the model-checking problem for the new logic is undecidable, which contrasts with the PSpace-complete complexity of the base logic. The proof strategy relies on reducing the emptiness problem for probabilistic automata (Madani, Hanks, and Condon 1999). Then, we extend with probabilities the knowing-how logic over LTS with indistinguishability classes, for which we have devised two cases. First (Sec. 4.2), we directly add probabilities to the logic presented in (Areces et al. 2021). In this case, that we call *non-adaptative*, the model-checking problem becomes undecidable contrasting the complexity of model checking the original logic, which is in PTime. This is proven by showing and using a variant of the result in (Madani, Hanks, and Condon 1999). The second proposal (Sec. 4.3), called here *adaptative*, is arguably suitable to model scenarios in which the agent has the ability to choose between one plan or another, among those that are considered equally good, depending on the particular situation. In fact, we compare expressivenes of all three logics with indistinguishability which helps to understand their utility. Also, for the adaptive case, we get that the model-checking problem is in PTime, another appealing characteristic of this logic. Along the paper, we discuss a running example to illustrate not only the behaviour of the logics, but also our design decisions.

2 Preliminaries

Words. Let Act be a finite set of *actions*, also called *alphabet* and let $\pi \in \mathsf{Act}^*$ be a word with alphabet Act. We use $|\pi|$ to denote the length of π (with $|\epsilon| = 0$) and $\pi[i]$, for $0 < i \leq |\pi|$, to denote the i-th element of π. For $0 \leq i \leq |\pi|$, $\pi[..i]$ denotes the i-th prefix of π, i.e., the initial segment of π up to (and including) the i-th position (with $\pi[..0] = \epsilon$). We say that π is a *prefix* of π', denoted $\pi \leq \pi'$, if $\pi = \pi'[..i]$ for some $0 \leq i \leq |\pi|$. For a set $\Pi \subseteq \mathsf{Act}^*$, let $\mathrm{pref}(\Pi) = \{\rho' \in \mathsf{Act}^* \mid \exists \rho \in \Pi \colon \rho' \leq \rho\}$ be the set of all prefixes of Π. We write $\mathrm{pref}(\pi)$ for $\mathrm{pref}(\{\pi\})$. Besides, $\pi\pi'$ denotes the concatenation of π followed by π'.

(Probabilistic) Labeled Transition Systems. A (discrete) *probability distribution* μ over a denumerable set S is a function $\mu : S \to [0, 1]$ such that $\mu(S) = \sum_{s \in S} \mu(s) = 1$. Let $\mathsf{Dist}(S)$ denote the set of all probability distributions on S, and let $\Delta_s \in \mathsf{Dist}(S)$ denote the Dirac distribution for $s \in S$, i.e., $\Delta_s(s) = 1$ and $\Delta_s(s') = 0$ for all $s' \in S$ such that $s' \neq s$.

Let Prop be a countable set of propositional symbols. A *probabilistic labeled transition System (PLTS)* (Segala 1995) is a tuple $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \to, \mathrm{V} \rangle$ where S is a finite set of *states*, Act is a finite set of *actions*, $\to \subseteq \mathrm{S} \times \mathsf{Act} \times \mathsf{Dist}(\mathrm{S})$ is the *(probabilistic) transition relation*, and $\mathrm{V} : \mathrm{S} \to \mathscr{P}(\mathsf{Prop})$ is the *valuation function*. We denote $s \xrightarrow{a} \mu$ whenever $(s, a, \mu) \in \to$ and let $T(s) = \{(a, \mu) \mid s \xrightarrow{a} \mu\}$ be the set of all transitions enabled in state s. For the case in which, for all $(s, a, \mu) \in \to$, μ is a Dirac distribution $\Delta_{s'}$ for some $s' \in \mathrm{S}$, we say that $\mathfrak{M}$ is a *labeled transition system (LTS)* and denote $s \xrightarrow{a} s'$ instead of $s \xrightarrow{a} \Delta_{s'}$.

An *execution* of $\mathfrak{M}$ is a finite or infinite alternating sequence of states and actions $s_0\, a_1\, s_1\, a_2\, s_2 \ldots$. $Exec_{\mathsf{f}} = \mathrm{S} \times (\mathsf{Act} \times \mathrm{S})^*$ denotes the set of all *finite executions* and $Exec_{\omega} = \mathrm{S} \times (\mathsf{Act} \times \mathrm{S})^{\omega}$ denotes the set of all *infinite executions*. We introduce the symbol $\perp$ to indicate that a finite execution has been intentionally ended, and let $CExec_{\mathsf{f}} = \mathrm{S} \times (\mathsf{Act} \times \mathrm{S})^* \times \{\perp\}$ denote the set of all *complete finite executions*. Infinite executions are also considered complete and hence $CExec = CExec_{\mathsf{f}} \cup Exec_{\omega}$ is the set of all *complete executions*.

For $\rho = s_0\, a_1\, s_1 \ldots s_{n-1}\, a_n\, s_n \in Exec_{\mathsf{f}}$ and $0 \leq k \leq n$, let $|\rho| = n$, $\mathrm{first}(\rho) = s_0$ and $\mathrm{last}(\rho) = s_n$. Let also $\rho[..k] = s_0\, a_1\, s_1 \ldots s_{k-1}\, a_k\, s_k$ be the k-th prefix of ρ. Similarly, we can define $\rho[..k]$ for $\rho \in CExec$. In particular, notice that $\rho[..0] = s_0$. We say that $\rho \in Exec_{\mathsf{f}} \cup CExec_{\mathsf{f}}$ is a *prefix* of $\rho' \in CExec$, denoted by $\rho \leq \rho'$, if $\rho = \rho'[..|\rho|]$ or $\rho = \rho'$ (this last case is needed if $\rho \in CExec_{\mathsf{f}}$). (The overloading of notation with respect to words should be harmless and easily understood by context.)

A *strategy* for a PLTS $\mathfrak{M}$ is a function $\sigma : Exec_{\mathsf{f}} \to \mathsf{Dist}((\mathsf{Act} \times \mathsf{Dist}(S)) \cup \{\bot\})$ that assigns a discrete probability distribution to each finite (non-complete) execution $\rho \in Exec_{\mathsf{f}}$ such that $\sigma(\rho)(a,\mu) > 0$ only if $\mathrm{last}(\rho) \xrightarrow{a} \mu$. Thus, a strategy can choose with some probability a valid transition after ρ or to intentionally terminate (in case $\sigma(\rho)(\bot) > 0$).

Let $\mathrm{Cyl}(\rho) = \{\rho' \in CExec \mid \rho \leq \rho'\}$ be the *cylinder set* induced by the finite execution $\rho \in Exec_{\mathsf{f}} \cup CExec_{\mathsf{f}}$. Notice that we only consider cylinders of complete executions and in particular $\mathrm{Cyl}(\rho) = \{\rho\}$ whenever $\rho \in CExec_{\mathsf{f}}$. A strategy σ and a state $s \in \mathrm{S}$ define a probability measure $\mathbb{P}^{\sigma}_{s}$ on the Borel sigma algebra generated by the set of all cylinder sets as follows. For $\rho = s_0\, a_1\, s_1 \ldots s_{n-1}\, a_n\, s_n \in Exec_{\mathsf{f}}$,

$$\mathbb{P}^{\sigma}_{s}(\mathrm{Cyl}(\rho)) = \tag{1}$$
$$\Delta_s(s_0) \cdot \prod_{i=1}^{n} \sum_{(a_i,\mu)\in T(s_{i-1})} \sigma(\rho[..(i-1)])(a_i,\mu) \cdot \mu(s_i)$$

and for $\rho = s_0\, a_1\, s_1 \ldots s_{n-1}\, a_n\, s_n \bot \in CExec_{\mathsf{f}}$,

$$\mathbb{P}^{\sigma}_{s}(\mathrm{Cyl}(\rho)) = \tag{2}$$
$$\Delta_s(s_0) \cdot \left(\prod_{i=1}^{n} \sum_{(a_i,\mu)\in T(s_{i-1})} \sigma(\rho[..(i-1)])(a_i,\mu) \cdot \mu(s_i) \right) \cdot\ \sigma(\rho)(\bot).$$

Carathedeory's extension theorem guarantees that $\mathbb{P}^{\sigma}_{s}$ is uniquely defined in the sigma algebra (Segala 1995).

Probabilistic Finite Automata. A PLTS is deterministic if $s \xrightarrow{a} \mu_1$ and $s \xrightarrow{a} \mu_2$ implies $\mu_1 = \mu_2$, for all $s \in \mathrm{S}$, $a \in \mathsf{Act}$, and $\mu_1, \mu_2 \in \mathsf{Dist}(\mathrm{S})$. A *probabilistic finite automata* (PFA) (Rabin 1963; Paz 1971) is a deterministic PLTS $\mathfrak{P} = \langle \mathrm{S}, \mathsf{Act}, \to, \mathrm{V}\rangle$ on the set $\mathsf{Prop} = \{\mathsf{init}, \mathsf{fin}\}$, i.e., $\mathrm{V} : \mathrm{S} \to \mathscr{P}(\{\mathsf{init}, \mathsf{fin}\})$, such that there is a single state $s_{\mathsf{i}} \in \mathrm{S}$, called *initial*, with $\mathsf{init} \in \mathrm{V}(s_{\mathsf{i}})$. $F = \{s \in \mathrm{S} \mid \mathsf{fin} \in \mathrm{V}(s)\}$ is the set of *final* or *accepting* states.

Given an execution $\rho = s_0\, a_1\, s_1 \ldots s_{n-1}\, a_n\, s_n \in Exec_{\mathsf{f}}$, let $\bar{\rho} = a_1\, a_2 \ldots a_n$. In particular $\bar{\rho} = \epsilon$ whenever $n = 0$. For a word $\pi \in \mathsf{Act}^*$, define the strategy σ_π such that for all $\rho \in Exec_{\mathsf{f}}$ with $\bar{\rho} \leq \pi$, (i) $\sigma_\pi(\rho)(a,\mu) = 1$ iff $\bar{\rho}a \leq \pi$ (in which case μ is unique), and (ii) $\sigma_\pi(\rho)(\bot) = 1$ iff either $\bar{\rho} = \pi$ or $\bar{\rho}\{a \mid \mathrm{last}(\rho) \xrightarrow{a} \mu\} \cap \mathrm{pref}(\pi) = \emptyset$. Since the PFA is deterministic, for any $s \in \mathrm{S}$, σ_π defines the same probability measure $\mathbb{P}^{\sigma_\pi}_{s}$ regardless of its definition for all ρ such that $\bar{\rho} \not\leq \pi$. The word π is accepted by the PFA $\mathfrak{P}$ with probability $q \in [0,1]$ if $\mathbb{P}^{\sigma_\pi}_{s_{\mathsf{i}}}(Succ(\pi, F)) \geq q$, where $Succ(\pi, F) = \{\rho\bot \in CExec_{\mathsf{f}} \mid \bar{\rho} = \pi \text{ and } \mathrm{last}(\rho) \in F\}$.

Proposition 1. *For a PFA and a bound $q \in [0,1]$, the following problems are undecidable:*

1. *determine if $\mathbb{P}^{\sigma_\pi}_{s_i}(Succ(\pi, F)) \geq q$ for some $\pi \in \mathsf{Act}^*$;*
2. *determine if $\mathbb{P}^{\sigma_\pi}_{s_i}(Succ(\pi, F)) < q$ for some $\pi \in \mathsf{Act}^*$.*

The first item corresponds to the emptiness problem in PFA (Paz 1971; Madani, Hanks, and Condon 1999). The second one, to the best of our knowledge, is new and can be proven from the first item using complete PFAs.

3 Knowing How Logics

We introduce first the common language we will use along this section. The set of formulas (a.k.a. the language) of $\mathcal{L}_{\mathsf{Kh}}$ is defined by the following BNF:

$$\varphi, \psi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid \mathsf{Kh}(\psi, \varphi),$$

where $p \in \mathsf{Prop}$. Other Boolean operators are defined as usual. The distinguished formula $\mathsf{Kh}(\psi, \varphi)$ should be read as *"the agent knows how to achieve φ given ψ"* meaning that there exists a plan that allows the agent to achieve the goal ψ whenever the assumed condition φ holds.

Linear Plans. In the most basic notion of knowing how –as defined in (Wang 2015; Wang 2018a; Wang 2018b)– a plan is a sequence of prescribed actions. Thus, formulas are interpreted over Labeled Transition Systems which indicate what actions are available for execution at each state and how they transform one state into another.

In order to determine when an agent knows how to achieve a goal, it is needed to characterize those plans that result appropriate for such a purpose. According to (Wang 2015), the notion of *strongly executable* provides such characterization and indicates that a plan is "fail proof". This notion was inspired by conformant planning (see e.g. (Cimatti, Roveri, and Traverso 1998)).

For an LTS $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \to, \mathrm{V}\rangle$, a *plan* is simply a word in Act^* (with ϵ being the empty plan). Given a plan $\pi \in \mathsf{Act}^*$, we define $\xrightarrow{\pi}$ as the composition $\xrightarrow{\pi[1]} \circ \ldots \circ \xrightarrow{\pi[|\pi|]}$. We say that a plan $\pi \in \mathsf{Act}^*$ is *strongly executable (SE)* at a state $s \in \mathrm{S}$ if and only if, for all $0 \leq i \leq |\pi| - 1$ and all $t \in \mathrm{S}$ such that $s \xrightarrow{\pi[..i]} t$, there is $v \in \mathrm{S}$ such that $t \xrightarrow{\pi[i+1]} v$. The plan π is SE at $A \subseteq \mathrm{S}$ if and only if it is SE at every $s \in A$. The notation $A \xrightarrow{\pi} G$ (for $A, G \subseteq \mathrm{S}$) indicates that for all $s \in A$, $s \xrightarrow{\pi} t$ implies $t \in G$.

Formulas are interpreted over pointed LTS, i.e., w.r.t. an LTS and a given state.

Definition 1. *Let $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \to, \mathrm{V}\rangle$ be an LTS and let $s \in \mathrm{S}$, the satisfiability relation $\models$ for $\mathcal{L}_{\mathsf{Kh}}$ is inductively defined as:*

$$\begin{array}{lll}
\mathfrak{M}, s \models p & \textit{iff}_{def} & p \in \mathrm{V}(s) \\
\mathfrak{M}, s \models \neg\varphi & \textit{iff}_{def} & \mathfrak{M}, s \not\models \varphi \\
\mathfrak{M}, s \models \psi \vee \varphi & \textit{iff}_{def} & \mathfrak{M}, s \models \psi \textit{ or } \mathfrak{M}, w \models \varphi \\
\mathfrak{M}, s \models \mathsf{Kh}(\psi, \varphi) & \textit{iff}_{def} & \textit{there is } \pi \in \mathsf{Act}^* \textit{ such that:} \\
 & & \quad (1)\ \pi \textit{ is SE at } [\![\psi]\!]^{\mathfrak{M}} \textit{ and} \\
 & & \quad (2)\ [\![\psi]\!]^{\mathfrak{M}} \xrightarrow{\pi} [\![\varphi]\!]^{\mathfrak{M}},
\end{array}$$

where: $[\![\chi]\!]^{\mathfrak{M}} = \{s \in \mathrm{S} \mid \mathfrak{M}, w \models \chi\}$.

The model-checking problem for a given logic is defined as follows, where models and formulas are instantiated with those corrresponding to each particular case.

Input: A model $\mathfrak{M}$, a state s in $\mathfrak{M}$ and a formula φ;
Output: True if $\mathfrak{M}, s \models \varphi$; False, otherwise.

Proposition 2 (Demri and Fervari 2023)**.** *The model-checking problem for $\mathcal{L}_{\mathsf{Kh}}$ over LTS is* PSpace*-complete.*

Indistinguishability Classes in Knowing How. We present the generalization of $\mathcal{L}_{\mathsf{Kh}}$ where the agent's perception is determined by an indistinguishability relation between plans (Areces et al. 2021; Areces et al. 2025), called herein $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$. W.l.o.g., we consider here a single agent, unlike the multi-agent presentation from previous works. Thus, the languages of $\mathcal{L}_{\mathsf{Kh}}$ and of $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$ coincide.

An *uncertainty-based* LTS (LTSU) is a tuple $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{U}, \mathrm{V}\rangle$ s.t. $\langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{V}\rangle$ is an LTS, and $\mathrm{U} \subseteq \mathscr{P}(\mathsf{Act}^*) \setminus \emptyset$ is a non-empty collection of pairwise disjoint non-empty sets of plans, i.e., (i) $\mathrm{U} \neq \emptyset$, (ii) for all $\Pi_1, \Pi_2 \in \mathrm{U}$, $\Pi_1 \neq \Pi_2$ implies $\Pi_1 \cap \Pi_2 = \emptyset$, and (iii) $\emptyset \notin \mathrm{U}$. The set U is called the *perception* of the agent.

Intuitively, $\mathrm{D} = \bigcup_{\Pi\in\mathrm{U}} \Pi$ is the set of plans that the agent is aware she has at her disposal, and each $\Pi \in \mathrm{U}$ is an indistinguishability class. Thus, U is in direct correspondence to an equivalence relation over D, where each $\Pi \in \mathrm{U}$ defines an equivalence class (see e.g. (Areces et al. 2025)). Notice that U is a partition of D representing the perception the agent has about the reality. Those plans in $\mathsf{Act}^* \setminus \mathrm{D}$ are not considered by the agent, even if they are suitable plans.

Given her indistinguishability over Act^*, the abilities of the agent do not depend on what a single plan can achieve, but rather on what a set of them can guarantee. Thus, for $\Pi \subseteq \mathsf{Act}^*$ and $A, G \subseteq \mathrm{S}$, we write $A \xrightarrow{\Pi} G$ if for all $s \in A$, $\pi \in \Pi$ and $t \in \mathrm{S}$, $s \xrightarrow{\pi} t$ implies $t \in G$.

In this new setting, we introduce the semantics of $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$.

Definition 2. *Let $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{U}, \mathrm{V}\rangle$ be an LTSU and let $s \in \mathrm{S}$, the satisfiability relation $\models$ for $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$ is defined inductively as usual for the Boolean operators and*

$$\mathfrak{M}, s \models \mathsf{Kh}(\psi, \varphi) \quad \textit{iff}_{def} \quad \textit{there is } \Pi \in \mathrm{U} \textit{ s.t. for all } \pi \in \Pi$$
$$(1)\ \pi \textit{ is SE at } [\![\psi]\!]^{\mathfrak{M}} \textit{ and}$$
$$(2)\ [\![\psi]\!]^{\mathfrak{M}} \xrightarrow{\pi} [\![\varphi]\!]^{\mathfrak{M}},$$

where: $[\![\chi]\!]^{\mathfrak{M}} = \{s \in \mathrm{S} \mid \mathfrak{M}, s \models \chi\}$.

In (Areces et al. 2021; Demri and Fervari 2023) two different instances of the model-checking problem for $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$ over LTSU were studied. In the former, U consists of a finite set of classes, where each of them is a finite set of plans. In the latter, U is finite but every class is characterized by a finite-state automaton, i.e., it is potentially infinite, subsuming the finite case. Both representations lead to the same result.

Proposition 3 (Demri and Fervari 2023)**.** *The model-checking problem for $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$ is in* PTime.

4 Probabilistic Knowing How

4.1 Linear Plans

Naturally, our first approach will be extending $\mathcal{L}_{\mathsf{Kh}}$ with some form of probabilistic behaviour. For this, we first need to understand how plans work in a probabilistic model. Thus, given a plan $\pi \in \mathsf{Act}^*$, we want to consider strategies that follow as faithfully as possible the plan π. This is captured in the next definition.

Definition 3. *A strategy σ is* π-compatible *if for all $\rho \in Exec_{\mathsf{f}}$ such that $\bar{\rho} \in \mathrm{pref}(\pi)$,*

1. *$\sigma(\rho)(a, \mu) > 0$ implies $\bar{\rho}a \in \mathrm{pref}(\pi)$, and*
2. *$\sigma(\rho)(\bot) > 0$ implies that either $\bar{\rho} = \pi$ or $\{\bar{\rho}a \mid \mathrm{last}(\rho) \xrightarrow{a} \mu\} \cap \mathrm{pref}(\pi) = \emptyset$.*

Let $Comp(\pi)$ denote the set of all π-compatible strategies.

The first item states that σ can chose an a-labeled transition after the partial plan $\bar{\rho}$ if the continuation of $\bar{\rho}a$ is also a partial plan. The second item states that σ is allowed to terminate after the partial plan $\bar{\rho}$ if either $\bar{\rho}$ is itself a valid plan or $\bar{\rho}$ cannot be continued by the PLTS within the plan.

It is important to remark that, since π is finite, for any π-compatible strategy σ and state s, $\mathbb{P}^{\sigma}_{s}(Exec_{\omega}) = 0$ (or equivalently, $\mathbb{P}^{\sigma}_{s}(CExec_{\mathsf{f}}) = 1$). That is, any π-compatible strategy leads to termination with probability 1. Also notice that in PFA, the strategy σ_π as defined in Sec. 2 is the only one π-compatible. By "only one", we mean that any strategy defined so that it satisfy the same conditions as σ_π yields the same probability measure $\mathbb{P}^{\sigma_\pi}_{s}$.

Example 1 (Running)**.** *Fig. 1 depicts the PLTS $\mathfrak{M}_{\mathrm{e}}$ modeling possible ways to escape a building in a fire emergency situation including alerting the event. There, states are represented by circles and distributions by the dot in the middle of transitions. The outgoing dashed arrows are labeled with the respective probability values. Thus, for instance, $s_0 \xrightarrow{lf} \mu_1$ with $\mu_1(s_1) = 0.2$ and $\mu_1(s_2) = 0.8$. Actions lf, st, and rm indicate that the exit might be reached through the lift, the stairs or the ramp. Actions pn and mb represent that the emergency is alerted through a panic button or by calling 911 using the mobile phone. Label ✓ indicates that the exit and the alert have been successfully performed. This happens only in states s_7, s_8, and s_{10} (thus $\mathrm{V}(s_7) = \mathrm{V}(s_8) = \mathrm{V}(s_{10}) = \{\checkmark\}$). States labeled with ✗ indicate that the last performed action has failed. In particular we distinguish the initial state by letting $\mathrm{V}(s_0) =$ init. Thus, it is possibe to take a lift through transition $s_0 \xrightarrow{lf} \mu_1$ and failed with probability 0.2 while exiting through the alternative lift (transition $s_0 \xrightarrow{lf} \mu_2$) fails with probability 1. Exiting through one of the stairs allows us to get to the panic button with probability 0.9 and enables the mobile call with probability 0.1, while exiting through the other stairs yields to the same situation but with the probabilities inverted. Exiting through the ramp allows us to press the panic button with probability 0.5 or to make the phone call also with probability 0.5. Notice that while the panic button always successfully rises the alarm, the mobile phone call may fail with probability 0.1.*

Let $\pi_1 = lf\ mb$. Define strategy σ_1 so that

$$\sigma_1(s_0)(lf, \mu_1) = \sigma_1(s_0)(lf, \mu_2) = 0.5$$
$$\sigma_1(s_0\ lf\ s_1)(\bot) = \sigma_1(s_0\ lf\ s_3)(\bot) = 1$$
$$\sigma_1(s_0\ lf\ s_2)(mb, \mu_6) = 1$$
$$\sigma_1(s_0\ lf\ s_2\ mb\ s_6)(\bot) = \sigma_1(s_0\ lf\ s_2\ mb\ s_7)(\bot) = 1.$$

It is easy to check that σ_1 is π_1-compatible. Also, note that σ'_1, defined so that $\sigma'_1(s_0)(lf, \mu_2) = \sigma'_1(s_0\ lf\ s_3)(\bot) = 1$, is also π_1-compatible despite that σ'_1 never manages to complete the plan π_1. Instead, σ''_1, defined so that $\sigma''_1(s_0)(lf, \mu_1) = \sigma'_1(s_0\ lf\ s_1)(\bot) = \sigma'_1(s_0\ lf\ s_2)(\bot) = 1$, is not π_1-compatible since $\sigma'_1(s_0\ lf\ s_2)(\bot) > 0$ but $s_2 \xrightarrow{mb} \mu_6$.

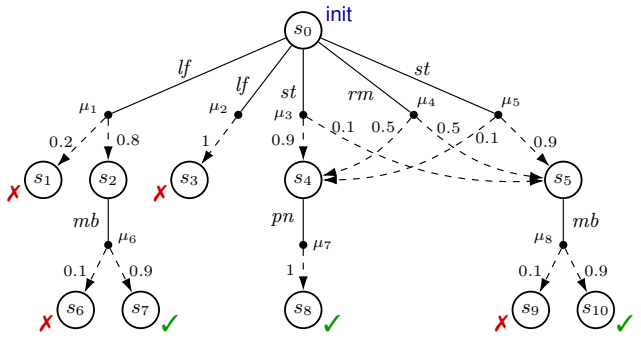

Figure 1: PLTS $\mathfrak{M}_{\mathrm{e}}$ modeling a fire emergency situation.

We define the set of *successful complete (finite) executions* that reach $G \subseteq \mathrm{S}$ following plan π by $Succ(\pi, G) = \{\rho\bot \in CExec_{\mathsf{f}} \mid \bar{\rho} = \pi \text{ and } \mathrm{last}(\rho) \in G\}$. We are interested in that any π-compatible strategy σ starting from a given state $s \in \mathrm{S}$ reaches a state in G with a minimum desired probability, say q. That is, we would like that

$$\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s}(Succ(\pi, G)) \geq q. \tag{3}$$

More generally, we would like that this holds from any particularly assumed state in a set $A \subseteq \mathrm{S}$ (say, a precondition). So, we write $A \xrightarrow{\pi}_{q} G$ if and only if for all state $s \in A$, the condition in eq. (3) holds. Thus $A \xrightarrow{\pi}_{q} G$ means that any state in A can reach the goal G with at least probability q following plan π.

Example 2. *Continuing with Ex. 1, let $G_{\checkmark}$ be the set of all states in which ✓ holds (i.e. $G_{\checkmark} = \{s_7, s_8, s_{10}\}$). Then $\{s_0\} \xrightarrow{\pi_1}_{q} G_{\checkmark}$ iff $q = 0$. This is a consequence of strategy σ_1' which is π_1-compatible but $\mathbb{P}^{\sigma_1'}_{s_0}(Succ(\pi_1, G_{\checkmark})) = 0$ (π_1 and σ_1' are as in Ex. 1).*

For plan $\pi_2 = st\ mb$, $\{s_0\} \xrightarrow{\pi_2}_{q} G_{\checkmark}$ iff $q \leq 0.09$. This is due to π_2-compatible strategy σ_2 defined so that

$$\sigma_2(s_0)(st, \mu_3) = 1$$
$$\sigma_2(s_0\ st\ s_4)(\bot) = \sigma_2(s_0\ st\ s_5)(mb, \mu_8) = 1$$
$$\sigma_2(s_0\ st\ s_5\ mb\ s_9)(\bot) = \sigma_2(s_0\ st\ s_5\ mb\ s_{10})(\bot) = 1,$$

with $\mathbb{P}^{\sigma_2}_{s_0}(Succ(\pi_2, G_{\checkmark})) = \mathbb{P}^{\sigma_2}_{s_0}(\{s_0\ st\ s_5\ mb\ s_{10}\bot\}) = 0.09$ (the first equality is a consequence of $s_0\ st\ s_5\ mb\ s_{10}\bot$ being the only complete execution in $Succ(\pi_2, G_{\checkmark})$ with non-zero probability). All other π_2-compatible strategy yields a probability larger than 0.09.

If we observe an LTS as a PLTS (as defined in Sec. 2) there is a strong connection between π-compatible strategies and strong executability of π. This connection is also lifted to relate $A \xrightarrow{\pi} G$ and $A \xrightarrow{\pi}_{1} G$ as stated in the next proposition.

Proposition 4. *For every LTS $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \to, \mathrm{V}\rangle$, state $s \in \mathrm{S}$, plan $\pi \in \mathsf{Act}^*$, and $A, G \subseteq \mathrm{S}$,*

1. *π is SE at s iff $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s}(Succ(\pi, \mathrm{S})) = 1$, and*
2. *π is SE at A and $A \xrightarrow{\pi} G$ iff $A \xrightarrow{\pi}_{1} G$.*

Proof sketch. Observe that $\mathbb{P}^{\sigma}_{s}(Succ(\pi, \mathrm{S})) = \mathbb{P}^{\sigma}_{s}(\{\rho\bot \mid \bar{\rho} = \pi \wedge \mathrm{first}(\rho) = s\})$ $(\star)$ for all $s \in \mathrm{S}$ and $\sigma \in Comp(\pi)$.

For $(\Rightarrow)$ of item 1, using $(\star)$ suppose by contradiction that that there exists $\sigma \in Comp(\pi)$ and $\rho \in Exec_{\mathsf{f}}$ such that $\bar{\rho} \neq \pi$, $\mathrm{first}(\rho) = s$, and $\mathbb{P}^{\sigma}_{s}(\{\rho\bot\}) > 0$. Analyzing $\bar{\rho} \neq \pi$ by cases $\bar{\rho} \lneq \pi$ (the prefix should be proper), $\pi \lneq \bar{\rho}$, and $\pi[i] \neq \bar{\rho}[i]$ for some $1 \leq i \leq \min(|\pi|, |\bar{\rho}|)$, the contradiction is obtained by eq. (2) and Def. 3.

For $(\Leftarrow)$ of item 1, we suppose $s \xrightarrow{\pi[..i]} t$ for $i < |\pi|$. Considering $(\star)$, eq. (2) and Def. 3, a π-compatible strategy σ that precisely follows $\pi[..i]$ from s to t (which exists by hypothesis), reveals the existence of a transition $t \xrightarrow{\pi[i+1]} v$.

For item 2, because of item 1, it suffices to show that for all $s \in A$ and SE plan π at s, $s \xrightarrow{\pi} t$ implies $t \in G$ iff $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s}(Succ(\pi, G)) = 1$. This follows by considering $(\star)$ and eqs. (1) and (2). □

Prop. 4 sets the bases to understand what we expect in a probabilistic setting: we would like to validate whether an agent knows how to achieve a goal *with at least some given probability of success*. That is, plans might not be perfect in a setting where actions are subject to failure or to random outcome. Therefore we introduce a probabilistic variant of Kh, denoted $\mathsf{Kh}^q(\psi, \varphi)$ here, which is interpreted as *"the agent knows how to achieve a goal φ given that ψ holds, with probability at least q"*.

Definition 4. *The language $\mathcal{L}_{\mathsf{Kh}^q}$ is defined by*

$$\varphi, \psi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid \mathsf{Kh}^q(\psi, \varphi),$$

where $p \in \mathsf{Prop}$ and $q \in (0, 1]$.

Item 2 in Prop. 4 shows that expression $A \xrightarrow{\pi}_{1} G$ captures both the idea of strong executability and the successful realization of a goal G. In a probabilistic setting, plans are allowed to fail, so the idea of strong executability is inconvenient. Instead, we are interested in that a plan π is realizable with at least probability q, and moreover, we are only interested in the realization of π if it achieves the goal G. That is why we define $A \xrightarrow{\pi}_{q} G$ to mean that from any state in A, the goal G is achieved with plan π with at least probability q. This is central for the definition of the semantics of $\mathcal{L}_{\mathsf{Kh}^q}$.

Definition 5. *Let $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \to, \mathrm{V}\rangle$ be a PLTS and let $s \in \mathrm{S}$. The satisfiability relation $\models$ for $\mathcal{L}_{\mathsf{Kh}^q}$ is defined inductively as usual for the Boolean operators and*

$$\mathfrak{M}, s \models \mathsf{Kh}^q(\psi, \varphi) \quad \textit{iff}_{\textit{def}} \quad \textit{there exists } \pi \in \mathsf{Act}^* \textit{ s.t. } [\![\psi]\!]^{\mathfrak{M}} \xrightarrow{\pi}_{q} [\![\varphi]\!]^{\mathfrak{M}},$$

where $[\![\chi]\!]^{\mathfrak{M}} = \{s \in \mathrm{S} \mid \mathfrak{M}, s \models \chi\}$.

In Def. 4 we requested that the probability bound is $q > 0$. The choice is due to the fact that $[\![\psi]\!]^{\mathfrak{M}} \xrightarrow{\pi}_{0} [\![\varphi]\!]^{\mathfrak{M}}$ is always true since any probability is larger than or equal to 0.

Example 3. *For the running example, for every state s, $\mathfrak{M}_{\mathrm{e}}, s \models \mathsf{Kh}^{0.5}(\mathsf{init}, \checkmark)$. This is explained through plan $rm\ pn$. Notice also that $\mathfrak{M}_{\mathrm{e}}, s \models \neg\mathsf{Kh}^{q}(\mathsf{init}, \checkmark)$ for any $q > 0.5$. One may think this is not the case since, for example, there is a way to successfully realize the plan $st\ pn$ with*

probability 0.9 *(starting with* $s_0 \xrightarrow{st} \mu_5$*). However, by starting instead with* $s_0 \xrightarrow{st} \mu_3$*, the same plan yields probability* 0.1 *(and hence* $[\![\mathsf{init}]\!]^{\mathfrak{M}} \xrightarrow{st\ pn}_q [\![\checkmark]\!]^{\mathfrak{M}}$ *iff* $q \leq 0.1$*).*

In the following we show that there is strong connection between $\mathcal{L}_{\mathsf{Kh}}$ and $\mathcal{L}_{\mathsf{Kh}^q}$ when limiting to LTSs models. We first present the next lemma which basically states that in LTS, plans that are realizable with some probability are also realized with probability 1.

Lemma 1. *For every LTS* $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{V} \rangle$*,* $s \in \mathrm{S}$ *and* $A, G \subseteq \mathrm{S}$*,*

1. $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s}(Succ(\pi, G)) > 0$ *iff* $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s}(Succ(\pi, G)) = 1$*, and*
2. $A \xrightarrow{\pi}_q G$ *for some* $q > 0$ *iff* $A \xrightarrow{\pi}_1 G$*.*

Proof. For implication ($\Rightarrow$) of item 1 suppose by contradiction that $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s}(Succ(\pi, G)) < 1$. Then there are $\sigma \in Comp(\pi)$ and $\rho \in Exec_{\mathsf{f}}$ such that $\mathbb{P}^{\sigma}_{s}(\{\rho\bot\}) > 0$ and $\rho\bot \notin Succ(\pi, G)$. Say $\rho = s_0\, a_1\, s_1 \ldots s_{n-1}\, a_n\, s_n$. Necessarily $s_0 = s$. Define strategy $\sigma^\star$ such that, for all $0 \leq i < n$, $\sigma^\star(s_0\, a_1\, s_1 \ldots s_{i-1}\, a_i\, s_i)(a_{i+1}, \Delta_{s_{i+1}}) = 1$, $\sigma^\star(\rho)(\bot) = 1$, and $\sigma^\star(\rho') = \sigma(\rho')$ for all other $\rho' \in Exec_{\mathsf{f}}$. Since σ is π-compatible and for all $0 \leq i < n$, $\sigma(s_0\, a_1\, s_1 \ldots s_{i-1}\, a_i\, s_i)(a_{i+1}, \Delta_{s_{i+1}}) > 0$, it should not be hard to check that $\sigma^\star$ is also π-compatible. Using eq. (2), we calculate that $\mathbb{P}^{\sigma^\star}_{s}(\{\rho\bot\}) = 1$ and hence $\mathbb{P}^{\sigma^\star}_{s}(Succ(\pi, G)) = 0$ contradicting the hypothesis that $\mathbb{P}^{\sigma^\star}_{s}(Succ(\pi, G)) > 0$. Since implication ($\Leftarrow$) is direct, item 1 is proved. Item 2 follows directly from item 1. □

Below we recursively define the mapping $\mathsf{rp} : \mathcal{L}_{\mathsf{Kh}^q} \to \mathcal{L}_{\mathsf{Kh}}$ which removes probability bounds from formulas by

$$\mathsf{rp}(p) = p \qquad \mathsf{rp}(\varphi \vee \psi) = \mathsf{rp}(\varphi) \vee \mathsf{rp}(\psi)$$
$$\mathsf{rp}(\neg\varphi) = \neg\mathsf{rp}(\varphi) \qquad \mathsf{rp}(\mathsf{Kh}^q(\psi, \varphi)) = \mathsf{Kh}(\mathsf{rp}(\psi), \mathsf{rp}(\varphi))$$

The following proposition states the exact correspondence between $\mathcal{L}_{\mathsf{Kh}}$ and $\mathcal{L}_{\mathsf{Kh}^q}$ over LTSs.

Proposition 5. *For all LTS* $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{V} \rangle$*,* $s \in \mathrm{S}$ *and* $\varphi \in \mathcal{L}_{\mathsf{Kh}^q}$*,* $\mathfrak{M}, s \models \varphi$ *iff* $\mathfrak{M}, s \models \mathsf{rp}(\varphi)$*.*

The proposition can be proven by structural induction on the formula, and using Lemma 1 and Prop. 4.

Now we proceed to analyze the computational behavior of $\mathcal{L}_{\mathsf{Kh}^q}$. Let $\mathfrak{M}$ be a PFA. Then $\mathfrak{M}, s \models \mathsf{Kh}^q(\mathsf{init}, \mathsf{fin})$ holds if there is a plan $\pi \in \mathsf{Act}^*$ such that $[\![\mathsf{init}]\!]^{\mathfrak{M}} \xrightarrow{\pi}_q [\![\mathsf{fin}]\!]^{\mathfrak{M}}$, that is, if $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s_{\mathsf{i}}}(Succ(\pi, F)) \geq q$, where $[\![\mathsf{init}]\!]^{\mathfrak{M}} = \{s_{\mathsf{i}}\}$ and $F = \{s \in \mathrm{S} \mid \mathsf{fin} \in \mathrm{V}(s)\}$. Since a PFA is deterministic, $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s_{\mathsf{i}}}(Succ(\pi, F)) = \mathbb{P}^{\sigma_\pi}_{s_{\mathsf{i}}}(Succ(\pi, F))$ with σ_π as in Sec. 2. Thus, checking $\mathfrak{M}, s \models \mathsf{Kh}^q(\mathsf{init}, \mathsf{fin})$ is equivalent to checking problem 1 in Prop. 1, yielding the following theorem.

Theorem 1. *The model-checking problem for* $\mathcal{L}_{\mathsf{Kh}^q}$ *is undecidable.*

The above result shows a huge jump in the computational behaviour of knowing-how: while model-checking for $\mathcal{L}_{\mathsf{Kh}}$ is PSpace-complete, considering probabilities leads to undecidability of the same problem.

4.2 Indistinguishable Classes

Just like we did for $\mathcal{L}_{\mathsf{Kh}}$, we want to extend $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$ to a probabilistic setting. A first natural approach is to keep the same spirit of $\mathcal{L}_{\mathsf{Kh}}$, in which plans are a priori commitments that are followed as pre-established.

First, we need to extend PLTSs with a perception U to reflect uncertainty.

Definition 6. *An* uncertainty-based LTS *(PLTSU) is a tuple* $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{U}, \mathrm{V} \rangle$ *such that* $\langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{V} \rangle$ *is a PLTS, and* $\mathrm{U} \subseteq \mathscr{P}(\mathsf{Act}^*) \setminus \emptyset$ *is the perception.*

In this setting, we reinterpret that $\mathsf{Kh}^q(\psi, \varphi)$ can only be successful if there is a class $\Pi \in \mathrm{U}$ such that every $\pi \in \Pi$ that starts at ψ ends in φ with probability at least q. Since the interpretation of the agent is that all plans in Π are equivalent, then all of them must perform as desired. This idea tries to extend to probability the non-probabilistic semantics of Def. 2. Under this new idea, we call the logic $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$, whose semantics is captured in the next definition.

Definition 7. *Let* $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{U}, \mathrm{V} \rangle$ *be a PLTSU and let* $s \in \mathrm{S}$*. The satisfiability relation* $\models$ *for* $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$ *is defined inductively as usual for the Boolean operators and*

$$\mathfrak{M}, s \models \mathsf{Kh}^q(\psi, \varphi) \quad \textit{iff}_{\textit{def}} \quad \textit{there exists } \Pi \in \mathrm{U} \textit{ s.t. for all } \pi \in \Pi, [\![\psi]\!]^{\mathfrak{M}} \xrightarrow{\pi}_q [\![\varphi]\!]^{\mathfrak{M}},$$

where $[\![\chi]\!]^{\mathfrak{M}} = \{s \in \mathrm{S} \mid \mathfrak{M}, s \models \chi\}$*.*

Example 4. *Let* $\pi_1 = lf\ mb$*,* $\pi_2 = st\ mb$*,* $\pi_3 = st\ pn$*,* $\pi_4 = rm\ mb$*, and* $\pi_5 = rm\ pn$*. Define* $\mathrm{U}_1 = \{\{\pi_1\}, \{\pi_2, \pi_3\}, \{\pi_4, \pi_5\}\}$ *and let* $\mathfrak{M}^1_{\mathrm{e}}$ *be the PLTSU extending the PLTS* $\mathfrak{M}_{\mathrm{e}}$ *with perception* U_1*. Then, for all* $s \in \mathrm{S}$*,* $\mathfrak{M}^1_{\mathrm{e}}, s \models \mathsf{Kh}^{0.45}(\mathsf{init}, \checkmark)$ *thanks to class* $\{\pi_4, \pi_5\}$*, since* $[\![\mathsf{init}]\!]^{\mathfrak{M}} \xrightarrow{\pi_4}_{0.45} [\![\checkmark]\!]^{\mathfrak{M}}$ *and* $[\![\mathsf{init}]\!]^{\mathfrak{M}} \xrightarrow{\pi_5}_{0.5} [\![\checkmark]\!]^{\mathfrak{M}}$*.*

Consider now $\mathrm{U}_2 = \{\{\pi_1\}, \{\pi_2, \pi_3, \pi_4, \pi_5\}\}$ *and let* $\mathfrak{M}^2_{\mathrm{e}}$ *be the PLTSU extending the PLTS* $\mathfrak{M}_{\mathrm{e}}$ *with perception* U_2*. Notice that* $\inf_{\sigma \in Comp(\pi_1)} \mathbb{P}^{\sigma}_{s_0}(Succ(\pi_1, \checkmark)) = 0$ *and hence class* $\{\pi_1\}$ *does not provide any probability of success. For class* $\{\pi_2, \pi_3, \pi_4, \pi_5\}$*, in particular,* $[\![\mathsf{init}]\!]^{\mathfrak{M}} \xrightarrow{\pi_2}_q [\![\checkmark]\!]^{\mathfrak{M}}$ *iff* $q \leq 0.09$*. Therefore,* $\mathfrak{M}^2_{\mathrm{e}}, s \models \neg\mathsf{Kh}^{0.1}(\mathsf{init}, \checkmark)$*.*

The following proposition states the exact correspondence between $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$ and $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$ over LTSUs.

Proposition 6. *For all LTSU* $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{U}, \mathrm{V} \rangle$*,* $s \in \mathrm{S}$ *and* $\varphi \in \mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$*,* $\mathfrak{M}, s \models \varphi$ *iff* $\mathfrak{M}, s \models \mathsf{rp}(\varphi)$*.*

The proposition can be proven by structural induction on the formula, and again using Lemma 1 and Prop. 4. It turns out that the model checking for $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$ is also undecidable contrary to the PTime problem to the relative non-probabilistic logic (Demri and Fervari 2023).

Theorem 2. *The model-checking problem for* $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$ *is undecidable.*

Proof. Let $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{U}, \mathrm{V} \rangle$ be a PLTSU so that $\mathrm{U} = \{\{\mathsf{Act}^*\}\}$ and $\langle \mathrm{S}, \mathsf{Act}, \rightarrow, \mathrm{V} \rangle$ is a complete PFA. (A PFA is complete if for all $s \in \mathrm{S}$ and $a \in \mathsf{Act}$ exists $\mu \in \mathsf{Dist}(\mathrm{S})$ s.t. $s \xrightarrow{a} \mu$. Also, any PFA has an equivalent complete PFA.)

$\mathfrak{M}, s \models \neg\mathsf{Kh}^q(\mathsf{init}, \mathsf{fin})$ holds if there is a plan $\pi \in \mathsf{Act}^*$ such that it does not happen that $[\![\mathsf{init}]\!]^{\mathfrak{M}} \xrightarrow{\pi}_q [\![\mathsf{fin}]\!]^{\mathfrak{M}}$,

or equivalently, if there is some $\pi \in \mathsf{Act}^*$ such that $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^{\sigma}_{s_\mathsf{i}}(Succ(\pi, F)) < q$, where $[\![\mathsf{init}]\!]^{\mathfrak{M}} = \{s_\mathsf{i}\}$ and $F = \{s \in \mathrm{S} \mid \mathsf{fin} \in \mathrm{V}(s)\}$. Since the PFA is deterministic, this is also equivalent to say that there is $\pi \in \mathsf{Act}^*$ such that $\mathbb{P}^{\sigma_\pi}_{s_\mathsf{i}}(Succ(\pi, F)) < q$. This is exactly problem 2 in Prop. 1, hence yielding the theorem. □

4.3 Indistinguishability with Adaptiveness

An alternative way to extend $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$ to a probabilistic setting is to consider *adaptive* plans. In reference to Ex. 4, notice that, although $\mathfrak{M}^1_{\mathrm{e}}, s \models \mathsf{Kh}^{0.45}(\mathsf{init}, \checkmark)$ (witnessed by class $\{\pi_4, \pi_5\}$), $\mathfrak{M}^1_{\mathrm{e}}, s \models \neg\mathsf{Kh}^{q}(\mathsf{init}, \checkmark)$ for all $q > 0.45$. However, since the agent's perception is that π_4 and π_5 are equivalent, there is no reason to commit to one plan that is about to fail while it is still possible to continue with the other. Since the class $\{\pi_4, \pi_5\}$ does not distinguish among its plans we might as well measure the class success as a whole. Thus, the decision whether to continue with π_4 and π_5 after transition $s_0 \xrightarrow{rm} \mu_4$ depends of the random outcome induced by μ_4. Notice that, by proceeding adaptively, the likelihood to succeed in this case reaches 0.95.

In this new view, the concept of a single plan compatibility becomes too strong. Thus, we rather ask strategies to be compatible with a whole class as follows.

Definition 8. *Given a set of plans $\Pi \subseteq \mathsf{Act}^*$, a strategy σ is* Π-compatible *if for all $\rho \in Exec_{\mathsf{f}}$ such that $\bar{\rho} \in \mathrm{pref}(\Pi)$,*

1. *$\sigma(\rho)(a, \mu) > 0$ implies $\bar{\rho}a \in \mathrm{pref}(\Pi)$, and*
2. *$\sigma(\rho)(\bot) > 0$ implies that either $\bar{\rho} \in \Pi$ or $\{\bar{\rho}a \mid \mathrm{last}(\rho) \xrightarrow{a} \mu\} \cap \mathrm{pref}(\Pi) = \emptyset$.*

Let $Comp(\Pi)$ denote the set of all Π-compatible strategies.

Notice that a strategy is $\{\pi\}$-compatible if and only if it is also π-compatible.

Example 5. *Consider perception U_1 of Ex. 4 and let $\Pi_{2,3} = \{\pi_2, \pi_3\}$ and $\Pi_{4,5} = \{\pi_4, \pi_5\}$ be two of its classes. Strategy σ_3, defined so that*

$$\begin{aligned}
&\sigma_3(s_0)(st, \mu_3) = \sigma_3(s_0)(st, \mu_5) = 0.5\\
&\sigma_3(s_0\ st\ s_4)(pn, \mu_7) = \sigma_3(s_0\ st\ s_5)(mb, \mu_8) = 1\\
&\sigma_3(s_0\ st\ s_4\ pn\ s_8)(\bot) = 1\\
&\sigma_3(s_0\ st\ s_5\ mb\ s_9)(\bot) = \sigma_3(s_0\ st\ s_5\ mb\ s_{10})(\bot) = 1,
\end{aligned}$$

is $\Pi_{2,3}$-compatible but not π_2-compatible or π_3-compatible. Notice the adaptive characteristic of σ_3 that chooses to perform $s_4 \xrightarrow{pn} \mu_7$ after st if in state s_4 but chooses $s_5 \xrightarrow{mb} \mu_8$ after st if in state s_5.

Similarly, strategy σ_4, discussed at the beginning of this subsection, is defined so that

$$\begin{aligned}
&\sigma_4(s_0)(rm, \mu_4) = 1\\
&\sigma_4(s_0\ rm\ s_4)(pn, \mu_7) = \sigma_4(s_0\ rm\ s_5)(mb, \mu_8) = 1\\
&\sigma_4(s_0\ rm\ s_4\ pn\ s_8)(\bot) = 1\\
&\sigma_4(s_0\ rm\ s_5\ mb\ s_9)(\bot) = \sigma_4(s_0\ rm\ s_5\ mb\ s_{10})(\bot) = 1,
\end{aligned}$$

and can be verified to be $\Pi_{4,5}$-compatible. However, it is neither π_4-compatible nor π_5-compatible.

We extend some concepts already defined for single plans to sets of plans. So, let $\Pi \subseteq \mathsf{Act}^*$ be a set of plans and let $G \subseteq \mathrm{S}$ be a set of goal states. The set of successful complete executions reaching G with a plan in Π is defined by $Succ(\Pi, G) = \{\rho\bot \in CExec_{\mathsf{f}} \mid \bar{\rho} \in \Pi \text{ and } \mathrm{last}(\rho) \in G\}$. The expression

$$\inf_{\sigma \in Comp(\Pi)} \mathbb{P}^{\sigma}_{s}(Succ(\Pi, G)) \geq q. \tag{4}$$

states that any Π-compatible strategy σ starting from a given state $s \in \mathrm{S}$ reaches a state in G with at least probability q. More generally, we extend this concept to a set of assumed starting states $A \subseteq \mathrm{S}$ by writing $A \xrightarrow{\Pi}_q G$ iff for all state $s \in A$, eq. (4) holds. Thus $A \xrightarrow{\Pi}_q G$ means that any state in A can reach the goal G with at least probability q following some plan in Π *in an adaptive manner*. Then, the agent can adapt to the best course of action in Π according to the available decisions along the chosen execution path.

Example 6. *Resuming the example, $\{s_0\} \xrightarrow{\Pi_{4,5}}_{0.95} G_{\checkmark}$ holds in $\mathfrak{M}^1_{\mathrm{e}}$. This is witnessed by strategy σ_4 and yields the value anticipated at the introduction of Sec. 4.3.*

In $\mathfrak{M}^2_{\mathrm{e}}$, if $\Pi_\star = \{\pi_2, \pi_3, \pi_4, \pi_5\} \in \mathrm{U}_2$, $\{s_0\} \xrightarrow{\Pi_\star}_{0.91} G_{\checkmark}$ witnessed by the $\Pi_\star$-compatible strategy σ_5 defined so that

$$\begin{aligned}
&\sigma_5(s_0)(st, \mu_3) = 1\\
&\sigma_5(s_0\ st\ s_4)(pn, \mu_7) = \sigma_5(s_0\ st\ s_5)(mb, \mu_8) = 1\\
&\sigma_5(s_0\ st\ s_4\ pn\ s_8)(\bot) = 1\\
&\sigma_5(s_0\ st\ s_5\ mb\ s_9)(\bot) = \sigma_5(s_0\ st\ s_5\ mb\ s_{10})(\bot) = 1.
\end{aligned}$$

For both $\mathfrak{M}^1_{\mathrm{e}}$ and $\mathfrak{M}^2_{\mathrm{e}}$, $\{s_0\} \xrightarrow{\{\pi_1\}}_q G_{\checkmark}$ iff $q = 0$, as expected from Ex. 2.

In this new context, $\mathsf{Kh}^q(\psi, \varphi)$ is reinterpreted so that there is an indistinguishable class $\Pi \in \mathrm{U}$ such that for every state that satisfies ψ, φ is reached with probability at least q following any plan in Π with every Π-compatible strategy. Under this new concept, we call the logic $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ and its semantics is captured in the next definition.

Definition 9. *Let $\mathfrak{M} = \langle \mathrm{S}, \mathsf{Act}, \to, \mathrm{U}, \mathrm{V}\rangle$ be a PLTSU and let $s \in \mathrm{S}$. The satisfiability relation $\models$ for $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ is defined inductively as usual for the Boolean operators and*

$$\mathfrak{M}, s \models \mathsf{Kh}^q(\psi, \varphi) \quad \textit{iff}_{\textit{def}} \quad \textit{there exists } \Pi \in \mathrm{U} \textit{ s.t. } [\![\psi]\!]^{\mathfrak{M}} \xrightarrow{\Pi}_q [\![\varphi]\!]^{\mathfrak{M}},$$

where $[\![\chi]\!]^{\mathfrak{M}} = \{s \in \mathrm{S} \mid \mathfrak{M}, s \models \chi\}$.

Notice that there is a *universal* cuantification on Π-compatible strategies implicit in the "inf" within expression $[\![\psi]\!]^{\mathfrak{M}} \xrightarrow{\Pi}_q [\![\varphi]\!]^{\mathfrak{M}}$ (see eq. (4)). This quantification parallels the universal quantification of plans in Π in Def. 2. However, in Def. 9 plans in Π are neither existential nor universally quantified. Instead they are *probabilistically* quantified for each Π-compatible strategy through expression $\mathbb{P}^{\sigma}_{s}(Succ(\Pi, G))$ within eq. (4).

Example 7. *For all $s \in \mathrm{S}$, $\mathfrak{M}^1_{\mathrm{e}}, s \models \mathsf{Kh}^{0.95}(\mathsf{init}, \checkmark)$ since $\{s_0\} \xrightarrow{\Pi_{4,5}}_{0.95} G_{\checkmark}$. However, $\mathfrak{M}^2_{\mathrm{e}}, s \models \neg\mathsf{Kh}^{0.95}(\mathsf{init}, \checkmark)$. This is a consequence of $\{s_0\} \xrightarrow{\{\pi_1\}}_q G_{\checkmark}$ iff $q = 0$ and $\{s_0\} \xrightarrow{\Pi_\star}_q G_{\checkmark}$ iff $q \leq 0.91$, the latter being witnessed by $\Pi_\star$-compatible strategy σ_5 (see Ex. 6).*

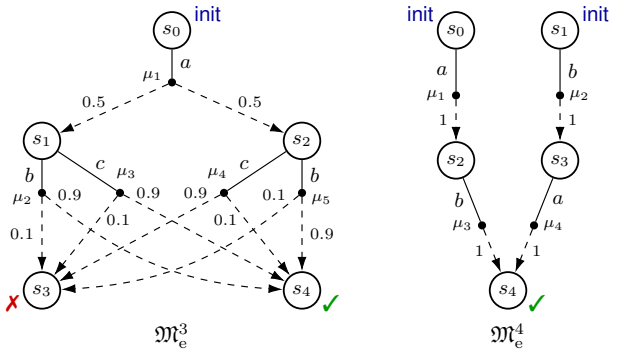

Figure 2: Counterexamples.

Logics $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$ and $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ are not directly related. On the one hand $\mathfrak{M}^1_{\mathrm{e}}, s \models_{\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}} \mathsf{Kh}^{0.95}(\mathsf{init}, \checkmark)$ (see Ex. 7) but $\mathfrak{M}^1_{\mathrm{e}}, s \models_{\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}} \neg\mathsf{Kh}^{0.95}(\mathsf{init}, \checkmark)$ (see Ex. 4). (Subscripts in $\models$ distinguish the logic in which the formula holds.) On the other hand, consider PLTSU $\mathfrak{M}^3_{\mathrm{e}}$ in Fig. 2 with $\mathrm{U}_3 = \{\{ab, ac\}\}$. It is not hard to check that $\mathfrak{M}^3_{\mathrm{e}}, s \models_{\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}} \mathsf{Kh}^{0.5}(\mathsf{init}, \checkmark)$ but $\mathfrak{M}^3_{\mathrm{e}}, s \models_{\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}} \neg\mathsf{Kh}^{0.5}(\mathsf{init}, \checkmark)$. The second one is explained by an $\{ab, ac\}$-compatible strategy that always chooses wrongly: at state s_1 it chooses the c transition while at state s_2 it chooses the b transition.

It is also the case that $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$ is not as strongly related to $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ as it is to $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$. Indeed, consider LTSU $\mathfrak{M}^4_{\mathrm{e}}$ in Fig. 2 with $\mathrm{U}_4 = \{\{ab, ba\}\}$. Then $\mathfrak{M}^4_{\mathrm{e}}, s \models_{\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}} \mathsf{Kh}^{1}(\mathsf{init}, \checkmark)$ but $\mathfrak{M}^4_{\mathrm{e}}, s \models_{\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}} \neg\mathsf{Kh}(\mathsf{init}, \checkmark)$. However, for an $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ formula φ in which Kh^q does not appear in the scope of a negation, if $\mathsf{rp}(\varphi)$ holds in an LTSU under $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}}$, then φ holds in the same LTSU under $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$. We show this after presenting the next lemma.

Lemma 2. *For every LTSU $\mathfrak{M}$ with perception* U, $s \in \mathrm{S}$, *and* $\Pi \in \mathrm{U}$,

1. *if* $\inf_{\sigma \in Comp(\pi)} \mathbb{P}^\sigma_s(Succ(\pi, G)) = 1$ *for all* $\pi \in \Pi$, *then* $\inf_{\sigma \in Comp(\Pi)} \mathbb{P}^\sigma_s(Succ(\Pi, G)) = 1$; *and*
2. *for all* $A, G \subseteq \mathrm{S}$ *if* $A \xrightarrow{\pi}_1 G$ *for all* $\pi \in \Pi$, $A \xrightarrow{\Pi}_1 G$.

Proof sketch. The proof of item 1 follows by assuming that there exist $\sigma \in Comp(\Pi)$ and $\rho \in Exec_{\mathsf{f}}$ with $\mathbb{P}^\sigma_s(\{\rho\bot\}) > 0$ and either $\bar\rho \notin \Pi$ or $\mathrm{last}(\rho) \notin G$ and deriving from σ a π-compatible strategy $\sigma^\star$ with $\pi \in \Pi$ such that $\mathbb{P}^{\sigma^\star}_s(Succ(\pi, G)) = 0$ contradicting the assumption. Item 2 follows directly from item 1. □

Proposition 7. *Let $\chi \in \mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ so that no operator Kh^q appears in the scope of $\neg$. For all LTSU $\mathfrak{M}$ and $s \in \mathrm{S}$, $\mathfrak{M}, s \models \mathsf{rp}(\chi)$ implies $\mathfrak{M}, s \models \chi$.*

Proof. The proof follows by distinguishing the cases in which χ contains some Kh^q operator or it does not. If χ does not contain Kh^q, then straightforwardly $\mathfrak{M}, s \models \mathsf{rp}(\chi)$ iff $\mathfrak{M}, s \models \chi$ by structural induction.

The case in which χ contains Kh^q also follows by structural induction. In particular, the case $\chi = \mathsf{Kh}^q(\psi, \varphi)$ follows from Lemma 2 (item 1) and Prop. 4 (item 2). □

In the rest of this section, we show that the problem of model checking for $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ is decidable provided that the classes of the agent's perception are regular languages. We first restate deterministic finite automata in our setting.

A *deterministic finite automaton (DFA)* is a PFA that is also an LTS. Let $\mathfrak{D}$ be a DFA. The language accepted by $\mathfrak{D}$ is defined by $L(\mathfrak{D}) = L(\mathfrak{D}, 1)$, i.e., it is the language accepted with probability 1 by $\mathfrak{D}$ seen as a PFS. We say that $\mathfrak{D}$ is *live* if for all $s \in \mathrm{S}$ there is some $\rho \in Exec_{\mathsf{f}}$ so that $\mathrm{first}(\rho) = s$ and $\mathrm{last}(\rho) \in F$ (F is the set of final states defined as for PFAs in Sec. 2). That is, $\mathfrak{D}$ is live if all of its states are involved in the acceptance of some word. Notice that for any regular language there is a live DFA that accepts it.

Let $\mathfrak{M} = \langle \mathrm{S}_\mathfrak{M}, \mathsf{Act}, \to_\mathfrak{M}, \mathrm{V}_\mathfrak{M}\rangle$ be a PLTS. Let $\mathfrak{D} = \langle \mathrm{S}_\mathfrak{D}, \mathsf{Act}, \to_\mathfrak{D}, \mathrm{V}_\mathfrak{D}\rangle$ be a live DFA with initial state t_{i} (i.e. $\mathsf{init} \in \mathrm{V}_\mathfrak{D}(t_{\mathsf{i}})$). The product PLTS of $\mathfrak{M}$ and $\mathfrak{D}$ is defined by $\mathfrak{M}\times\mathfrak{D} = \langle \mathrm{S}_\mathfrak{M} \times \mathrm{S}_\mathfrak{D}, \mathsf{Act}, \to, \mathrm{V}\rangle$ where $\mathrm{V}(s,t) = \mathrm{V}_\mathfrak{M}(s) \cup \mathrm{V}_\mathfrak{D}(t)$ and $\to \in (\mathrm{S}_\mathfrak{M} \times \mathrm{S}_\mathfrak{D}) \times \mathsf{Act} \times \mathsf{Dist}(\mathrm{S}_\mathfrak{M} \times \mathrm{S}_\mathfrak{D})$ is the smallest relation such that

$$s \xrightarrow{a}_\mathfrak{M} \mu \text{ and } t \xrightarrow{a}_\mathfrak{D} \Delta_{t'} \text{ imply } (s,t) \xrightarrow{a} \mu\otimes\Delta_{t'},$$

with $\mu\otimes\Delta_{t'}$ being the usual product of distributions defined by $\mu\otimes\Delta_{t'}(s,t) = \mu(s)\cdot\Delta_{t'}(t)$.

The following lemma is central to provide a model checking algorithm for $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$.

Lemma 3. *Let $\mathfrak{M}$ be a PLTS and let $G \subseteq \mathrm{S}_\mathfrak{M}$ be a set of goal states. Let $\Pi \subseteq \mathsf{Act}^*$ be a regular language accepted by the live DFA $\mathfrak{D}$ with initial state $t_i \in \mathrm{S}_\mathfrak{D}$. Define $R_{G\times F} = \{\rho \in CExec^{\mathfrak{M}\times\mathfrak{D}} \mid \mathrm{last}(\rho[..i]) \in G\times F$ for some $i \geq 0\}$ (viz., the set of complete executions that reach simultaneously both the goal and a final accepting state). Besides, let $s_i \in \mathrm{S}_\mathfrak{M}$. Then:*

1. *For all strategy σ of $\mathfrak{M}\times\mathfrak{D}$, there is a Π-compatible strategy $\sigma^\star$ of $\mathfrak{M}$ s.t. $\mathbb{P}^\sigma_{(s_i,t_i)}(R_{G\times F}) = \mathbb{P}^{\sigma^\star}_{s_i}(Succ(\Pi, G))$.*
2. *For all Π-compatible strategy σ of $\mathfrak{M}$ there is a strategy $\sigma^\star$ of $\mathfrak{M}\times\mathfrak{D}$ s.t. $\mathbb{P}^\sigma_{s_i}(Succ(\Pi, G)) = \mathbb{P}^{\sigma^\star}_{(s_i,t_i)}(R_{G\times F})$.*
3. $\inf_{\sigma\in Comp(\Pi)} \mathbb{P}^\sigma_{s_i}(Succ(\Pi, G)) = \inf_\sigma \mathbb{P}^\sigma_{(s_i,t_i)}(R_{G\times F})$.
4. *For $A \subseteq \mathrm{S}_\mathfrak{M}$, $A \xrightarrow{\Pi}_q G$ iff $\inf_{s\in A} \inf_\sigma \mathbb{P}^\sigma_{(s_i,t_i)}(R_{G\times F}) \geq q$.*

Proof. We first state some general facts that would later simplify the proof of each of the items of the lemma. First it is not hard to verify that

$$R_{G\times F} = \bigcup_{\rho\in\mathcal{R}} \mathrm{Cyl}(\rho) \qquad \text{and} \tag{5}$$

$$Succ(\Pi, G) = \bigcup_{\rho\in\mathcal{S}} \mathrm{Cyl}(\rho) \tag{6}$$

where

$$\mathcal{R} = \{\rho \in Exec^{\mathfrak{M}\times\mathfrak{D}}_{\mathsf{f}} \mid \mathrm{last}(\rho) \in G\times F \text{ and for all } 0 \leq k < |\rho|, \mathrm{last}(\rho[..k]) \notin G\times F\}, \text{ and}$$

$$\mathcal{S} = \{\rho \in Exec^{\mathfrak{M}}_{\mathsf{f}} \mid \bar\rho \in \Pi,\ \mathrm{last}(\rho) \in G \text{ and for all } 0 \leq k < |\rho|, \text{ either } \mathrm{last}(\rho[..k]) \notin G \text{ or } \bar\rho[..k] \notin \Pi\}.$$

Let $\rho, \rho' \in \mathcal{R}$. Then, if $\rho \leq \rho'$, necessarily $\rho = \rho'$. Therefore, if $\rho \neq \rho'$, $\mathrm{Cyl}(\rho) \cap \mathrm{Cyl}(\rho') = \emptyset$. Also notice that if $\rho \in \mathcal{R}$, $\bar{\rho}$ is accepted by $\mathfrak{D}$ and hence $\bar{\rho} \in \Pi$. Similarly, for $\rho, \rho' \in \mathcal{S}$, if $\rho \leq \rho'$ then $\rho = \rho'$. Thus, if instead $\rho \neq \rho'$, $\mathrm{Cyl}(\rho) \cap \mathrm{Cyl}(\rho') = \emptyset$.

Let $\mathcal{R}_{\mathsf{i}} = \{\rho \in \mathcal{R} \mid \mathrm{first}(\rho) = (s_{\mathsf{i}}, t_{\mathsf{i}})\}$ and $\mathcal{S}_{\mathsf{i}} = \{\rho \in \mathcal{S} \mid \mathrm{first}(\rho) = s_{\mathsf{i}}\}$. By eq. (5), for every strategy σ,

$$\mathbb{P}^{\sigma}_{(s_{\mathsf{i}},t_{\mathsf{i}})}(R_{G\times F}) = \textstyle\sum_{\rho\in\mathcal{R}_{\mathsf{i}}} \mathbb{P}^{\sigma}_{(s_{\mathsf{i}},t_{\mathsf{i}})}(\mathrm{Cyl}(\rho)). \tag{7}$$

Similarly, by eq. (6), for every strategy σ,

$$\mathbb{P}^{\sigma}_{s_{\mathsf{i}}}(Succ(\Pi, G)) = \textstyle\sum_{\rho\in\mathcal{S}_{\mathsf{i}}} \mathbb{P}^{\sigma}_{s_{\mathsf{i}}}(\mathrm{Cyl}(\rho)). \tag{8}$$

Let $\rho = (s_0, t_0)\, a_1\, (s_1, t_1) \ldots (s_{n-1}, t_{n-1})\, a_n\, (s_n, t_n) \in Exec^{\mathfrak{M}\times\mathfrak{D}}_{\mathsf{f}}$. Define the mapping proj by $\mathrm{proj}(\rho) = s_0\, a_1\, s_1 \ldots s_{n-1}\, a_n\, s_n$. It turns out that proj is a bijection from $\mathcal{R}_{\mathsf{i}}$ to $\mathcal{S}_{\mathsf{i}}$. Indeed, suppose $\rho, \rho' \in \mathcal{R}_{\mathsf{i}}$ and ρ is as before. Then $\rho' = (s_0, t'_0)\, a_1\, (s_1, t'_1) \ldots (s_{n-1}, t'_{n-1})\, a_n\, (s_n, t'_n)$. Notice that $(s_0, t_0) = (s_{\mathsf{i}}, t_{\mathsf{i}}) = (s_0, t'_0)$. Since $\mathfrak{D}$ is deterministic, then necessarily $t_k = t'_k$ for all $0 \leq k \leq n$, which shows that proj is injective. To prove that proj is surjective, let $\rho = s_0\, a_1\, s_1 \ldots s_{n-1}\, a_n\, s_n \in \mathcal{S}_{\mathsf{i}}$. Then $s_0 = s_{\mathsf{i}}$ and, since $\bar{\rho} \in \Pi$, $a_1\, a_2 \ldots a_n$ is accepted by $\mathfrak{D}$. As a consequence there are $t_0, t_1, \ldots, t_n \in \mathrm{S}_{\mathfrak{D}}$ with $t_0 = t_{\mathsf{i}}$, $t_n \in F$ and $t_k \xrightarrow{a_{k+1}} \Delta_{t_{k+1}}$ for all $0 \leq k < n$. Hence, $\rho' = (s_0, t_0)\, a_1\, (s_1, t_1) \ldots (s_{n-1}, t_{n-1})\, a_n\, (s_n, t_n) \in Exec^{\mathfrak{M}\times\mathfrak{D}}_{\mathsf{f}}$. If, for some $k < n$, $(s_k, t_k) \in G\times F$, $a_1\, a_2 \ldots a_k \in \Pi$ and $s_k \in G$ contradicting that $\rho \in \mathcal{S}_{\mathsf{i}}$. Therefore $\rho' \in \mathcal{R}_{\mathsf{i}}$.

Now we address item 1. Let σ be a strategy for $\mathfrak{M}\times\mathfrak{D}$. We define strategy $\sigma^\star$ for $\mathfrak{M}$ as follows. Let $\rho = (s_0, t_0)\, a_1\, (s_1, t_1) \ldots (s_{n-1}, t_{n-1})\, a_n\, (s_n, t_n) \in \mathcal{R}_{\mathsf{i}}$ and $0 \leq k \leq |\rho|$, then:

1. If $k = |\rho|$, let $\sigma^\star(\mathrm{proj}(\rho))(\bot) = 1$. Since $\overline{\mathrm{proj}(\rho)} = \bar{\rho} \in \Pi$, then $\sigma^\star$ satisfies item 2 of Def. 8 for $\mathrm{proj}(\rho)$.
2. If instead $k < |\rho|$, let $\sigma^\star(\mathrm{proj}(\rho)[..k])(a_{k+1}, \mu) = \sigma(\rho[..k])(a_{k+1}, \mu \otimes \Delta_{t_{k+1}})$ for all $\mu \in \mathsf{Dist}(\mathrm{S})$. Since $\overline{\mathrm{proj}(\rho)[..k]} \in \mathrm{pref}(\Pi)$ and $\overline{\mathrm{proj}(\rho)[..k]}a_{k+1} \in \mathrm{pref}(\Pi)$, then $\sigma^\star$ satisfies item 1 of Def. 8 for $\mathrm{proj}(\rho)[..k]$.

Since proj is a bijection from $\mathcal{R}_{\mathsf{i}}$ to $\mathcal{S}_{\mathsf{i}}$, $\sigma^\star$ is well defined for all executions that are prefixes of some execution in $\mathcal{S}_{\mathsf{i}}$ and, moreover, it satisfies the conditions of Def. 8 for all those executions. Then $\sigma^\star$ can be defined conveniently for any other execution so that it is indeed Π-compatible.

Following eqs. (1) and (2), it should not be hard to see that $\mathbb{P}^{\sigma}_{(s_{\mathsf{i}},t_{\mathsf{i}})}(\mathrm{Cyl}(\rho)) = \mathbb{P}^{\sigma^\star}_{s_{\mathsf{i}}}(\{\mathrm{proj}(\rho)\bot\}) = \mathbb{P}^{\sigma^\star}_{s_{\mathsf{i}}}(\mathrm{Cyl}(\mathrm{proj}(\rho)))$. Hence, from eqs. (7) and (8), item 1 follows.

The proof of item 2 follows in a similar way to that of item 1. Item 3 follows from items 1 and 2, and item 4 is a direct consequence of item 3. □

As a consequence of item 4 in Lemma 3, the validity of $A \xrightarrow{\Pi}_q G$ can be checked on a simple extension of the product PLTS $\mathfrak{M}\times\mathfrak{D}$. More precisely, extend $\mathfrak{M}\times\mathfrak{D}$ with a fresh state s_A, a new fresh action $\tau \notin \mathsf{Act}$ and a transition $s_A \xrightarrow{\tau} \Delta_s$ for each $s \in A$. Then $A \xrightarrow{\Pi}_q G$ iff $\inf_\sigma \mathbb{P}^{\sigma}_{s_A}(R_{G\times F}) \geq q$. The full model checking algorithm is implemented recursively as given in Alg. 1 where function MeanReachProb($G\times F_i$) (with F_i being the set of final states of $\mathfrak{D}_i$) calculates the minimun reachability probability $\inf_\sigma \mathbb{P}^{\sigma}_{s_A}(R_{G\times F_i})$ using standard techniques (Puterman 1994; Baier and Katoen 2008). In particular, notice that in line 11, the function returns the whole set S of states if the i-th DFA witnesses the validity of $\mathsf{Kh}^q(\psi, \chi)$ and, in 12, it returns the empty set if no DFA witnesses $\mathsf{Kh}^q(\psi, \chi)$.

Algorithm 1 Model checking algorithm for $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$

Input: PLTSU $\mathfrak{M}$, with perception given as a family of live DFA $\{\mathfrak{D}_i\}_{i\in I}$, $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ formula φ
Output: $[\![\varphi]\!]^{\mathfrak{M}}$
Function: Check(φ)

```
1: switch (φ) do
2:   case p ∈ Prop: return {s ∈ S | p ∈ V(s)}
3:   case ¬ψ: return S \ Check(ψ)
4:   case ψ ∧ χ: return Check(ψ) ∩ Check(χ)
5:   case Kh^q(ψ, χ):
6:     A := Check(ψ)
7:     G := Check(χ)
8:     for each i ∈ I do
9:       Construct M×D_i extended with s_A
10:      qmin := MeanReachProb(G×F_i)
11:      if (qmin ≥ q) then return S
12:    return ∅
```

Since $\inf_\sigma \mathbb{P}^{\sigma}_{s_A}(R_{G\times F_i})$ can be formulated as a linear optimization problem (Puterman 1994; Bianco and de Alfaro 1995; Baier and Katoen 2008), MeanReachProb($G\times F_i$) can be implemented as a polynomial time algorithm (e.g. (Karmarkar 1984)). Thus, a simple inspection of Alg. 1 yields the next theorem.

Theorem 3. *The model-checking problem for $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$ is decidable provided each class in* U *is a regular language. Moreover, if each* $\Pi \in$ U *is given as input as a live DFA, the problem is in* PTime.

5 Final Remarks

We investigated the model-checking problem for probabilistic variants of knowing-how logics. In particular, we proved that for the direct extensions of the works from (Wang 2015) and (Areces et al. 2021) the problem is undecidable. Then, we detect an interesting variant in which the agent proceeds adaptatively, for which model checking is decidable in polynomial time. Arguably, our "adaptative" view subtly differs from the "close-loop" policy in control, since our adaptative agent chooses with no criteria among the decisions she considers equivalent. Furthermore, our results shed some new light for understanding constrained knowing how logics, and introduce novel formalisms with interesting applications.

For future work, it would be interesting to characterize the exact expressive power of $\mathcal{L}^{\mathrm{U}}_{\mathsf{Kh}^q}$ and $\mathcal{L}^{\mathrm{A}}_{\mathsf{Kh}^q}$, and compare them. In particular, define associated notions of bisimulation and prove characterization theorems. Also, it would be interesting to define proof systems for the logics we investigated. Finally, our next step will be to implement Alg. 1 into the PRISM tool (Kwiatkowska, Norman, and Parker 2011).

Acknowledgments

This work was supported by Agencia I+D+i grant PICT 2021-00400, the EU H2020 research and innovation programme under the Marie Skłodowska-Curie grant agreements 101008233 (MISSION), the IRP SINFIN, SeCyT-UNC grants 33620230100384CB (MECANO) and 33620230100178CB, and as part of France 2030 program ANR-11-IDEX-0003.

References

Aminof, B.; Kwiatkowska, M.; Maubert, B.; Murano, A.; and Rubin, S. 2019. Probabilistic strategy logic. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, 32–38. International Joint Conferences on Artificial Intelligence Organization.

Areces, C.; Fervari, R.; Saravia, A. R.; and Velázquez-Quesada, F. R. 2021. Uncertainty-based semantics for multi-agent knowing how logics. In *18th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2021)*, volume 335 of *EPTCS*, 23–37. Open Publishing Association.

Areces, C.; Fervari, R.; Saravia, A. R.; and Velázquez-Quesada, F. R. 2025. Uncertainty-based knowing how logic. *Journal of Logic and Computation* 35(1):1–35.

Baier, C., and Katoen, J. 2008. *Principles of model checking*. MIT Press.

Baier, C.; Forejt, V.; de Alfaro, L.; and Kwiatkowska, M. 2018. Model checking probabilistic systems. In Clarke, E. M.; Henzinger, T. A.; Veith, H.; and Bloem, R., eds., *Handbook of Model Checking*. Springer. 963–999.

Berthon, R.; Katoen, J.; Mittelmann, M.; and Murano, A. 2024. Natural strategic ability in stochastic multi-agent systems. In *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024*, 17308–17316. AAAI Press.

Bianco, A., and de Alfaro, L. 1995. Model checking of probabilistic and nondeterministic systems. In *Foundations of Software Technology and Theoretical Computer Science*, 499–513. Berlin, Heidelberg: Springer Berlin Heidelberg.

Bulling, N., and Jamroga, W. 2009. What agents can probably enforce. *Fundamenta Informaticae* 93(1-3):81–96.

Chen, T., and Lu, J. 2007. Probabilistic alternating-time temporal logic and model checking algorithm. In *Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery - Volume 02*, FSKD '07, 35–39. USA: IEEE Computer Society.

Cimatti, A.; Roveri, M.; and Traverso, P. 1998. Strong planning in non-deterministic domains via model checking. In *Proceedings of the Fourth International Conference on Artificial Intelligence Planning Systems, Pittsburgh, Pennsylvania, USA, 1998*, 36–43. AAAI.

Demri, S., and Fervari, R. 2023. Model-checking for ability-based logics with constrained plans. In *37th AAAI Conference on Artificial Intelligence (AAAI 2023)*, 6305–6312. AAAI Press.

Fervari, R.; Herzig, A.; Li, Y.; and Wang, Y. 2017. Strategically knowing how. In *26th International Joint Conference on Artificial Intelligence (IJCAI 2017)*, 1031–1038. International Joint Conferences on Artificial Intelligence.

Herzig, A., and Troquard, N. 2006. Knowing how to play: uniform choices in logics of agency. In *5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006)*, 209–216. ACM.

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

Hintikka, J. 1962. *Knowledge and Belief*. Cornell University Press.

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

Karmarkar, N. 1984. A new polynomial-time algorithm for linear programming. *Comb.* 4(4):373–396.

Kushmerick, N.; Hanks, S.; and Weld, D. S. 1995. An algorithm for probabilistic planning. *Artificial Intelligence* 76(1):239–286. Planning and Scheduling.

Kwiatkowska, M. Z.; Norman, G.; and Parker, D. 2011. PRISM 4.0: Verification of probabilistic real-time systems. In *Computer Aided Verification - 23rd International Conference, CAV 2011, Snowbird, UT, USA, July 14-20, 2011. Proceedings*, volume 6806 of *Lecture Notes in Computer Science*, 585–591. Springer.

Lespérance, Y.; Levesque, H. J.; Lin, F.; and Scherl, R. B. 2000. Ability and knowing how in the situation calculus. *Studia Logica* 66(1):165–186.

Li, Y., and Wang, Y. 2017. Achieving while maintaining: A logic of knowing how with intermediate constraints. In *7th Indian Conference on Logic and Its Applications (ICLA 2017)*, LNCS, 154–167. Springer.

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

Madani, O.; Hanks, S.; and Condon, A. 1999. On the undecidability of probabilistic planning and infinite-horizon partially observable markov decision problems. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence and Eleventh Conference on Innovative Applications of Artificial Intelligence*, 541–548. AAAI Press / The MIT Press.

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

Moore, R. 1985. A formal theory of knowledge and action. In *Formal Theories of the Commonsense World*. Ablex Publishing Corporation.

Naumov, P., and Tao, J. 2018. Together we know how to achieve: An epistemic logic of know-how. *Artificial Intelligence* 262:279–300.

Naumov, P., and Tao, J. 2019. Knowing-how under uncertainty. *Artificial Intelligence* 276:41–56.

Paz, A. 1971. *Introduction to probabilistic automata*. Computer science and applied mathematics. Academic Press, Inc.

Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Statistics. Wiley.

Rabin, M. O. 1963. Probabilistic automata. *Inf. Control.* 6(3):230–245.

Russell, S., and Norvig, P. 2020. *Artificial Intelligence: A Modern Approach*. Pearson, 4th edition.

Segala, R. 1995. *Modeling and verification of randomized distributed real-time systems*. Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA.

van der Hoek, W.; van Linder, B.; and Meyer, J. C. 2000. On agents that have the ability to choose. *Stud Logica* 66(1):79–119.

Wang, Y. 2015. A logic of knowing how. In *5th International Workshop on Logic, Rationality, and Interaction (LORI 2015)*, LNCS, 392–405. Springer.

Wang, Y. 2018a. Beyond knowing that: a new generation of epistemic logics. In *J. Hintikka on knowledge and game theoretical semantics*. Springer. 499–533.

Wang, Y. 2018b. A logic of goal-directed knowing how. *Synthese* 195(10):4419–4439.