Knowing what to do

Li, Yanjun

Publication date:
2017

Yanjun Li

Knowing What to Do

A logical approach to planning and knowing how

This research was funded by China Scholarship Council (CSC).

ISBN: 978-94-034-0033-4
ISBN: 978-94-034-0034-1 (eBook)

university of groningen

Knowing What to Do

A logical approach to planning and knowing how

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. E. Sterken
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Thursday 21 September 2017 at 11.00 hours

by

Yanjun Li

born on 18 February 1985
in Hebei, China

Supervisor
Prof. B.P. Kooi

Co-supervisor
Dr. J.A. van Laar

Assessment Committee
Prof. L.C. Verbrugge
Prof. V. Goranko
Prof. H.P. van Ditmarsch

Contents

Acknowledgements

First of all, I am in debt to Barteld Kooi, my supervisor. In the past four years, your guidance and encouragement helped me a lot in my research. You spent lots of time reading drafts of my papers, gave me thoughtful comments and detailed suggestions. Moreover, you are not only a great scholar, a supportive supervisor, but also a cheerful and warm-hearted person. Your help makes my Groningen life much easier.

I would like to thank Jan Albert van Laar, my co-supervisor. Thank you for your helpful comments. Especially, thank you for carefully reading the manuscript and helping me to improve the language of this thesis.

I am grateful to Yanjing Wang. You are like another supervisor to me. It is you that encourage me coming to Groningen and pursuing a second PhD, and it is your help that enables me to complete this thesis.

I am thankful to Beihai Zhou who is the supervisor of my first PhD at Peking University. This second degree would not have happened without your support. I am thankful to Yanjing Zhang at Hebei University. Thanks for all your help.

I also would like to thank my co-authors: Barteld Kooi, Yanjing Wang, Raul Fervari and Andreas Herzig for their genuine contributions. I am also grateful to the anonymous referees who reviewed the papers that compose the main body of this thesis.

I thank the respectful members of my assessment committee: Rineke Verbrugge, Valentin Goranko and Hans van Ditmarsch. Thank you for your time and your valuable comments.

I would like to thank all my colleagues in Groningen. You make the Faculty of Philosophy a pleasant place to work. For their helpful comments, I would like to thank the members of PCCP and the members of WiP, including but not limited to, Job de Grefte, Coos Engelsma, Benjamin Bewersdorf, Stipe Pandzic, Yuri Santos, Max Bialek, Louwe B. Kuijer, Wessel van Dommelen, Rohan French, Li-Chih Lin, Sanne Hupkes, Sander Verhaegh, Franziska Köder, Leon Geerdink, Oberto Marrama, Sjoerd Griffioen, Han Thomas Adriaenssen, Paulien Talsma-Snellen, Bianca Bosman. For their help with technical and financial issues, I would like to thank Marga Hids, Anita Willems-Veenstra, Siepie Blom, Fré Morrees, Hauke de Vries, and Eric Ruben.

I would like to thank all my friends in Groningen: Bin Jiang, Qian Li, Linyang Li, Peng Wang, Yu Sun, Yingruo Wang, Jianjun Tang & Wangli Cheng, Bo Zhang & Tingting Qi, Yuanze Wang & Jingjing Deng, Xiaoming Zhang & Ping Xu, Jiquan Wu & Zhonghong Ma, Peiliang Zhao & Lijuan Xing, Qingkai Yang & Yuhan Wang, Shengjie Xiao & Baiyu An, Yanglei Yi, Haojie Cao, Jing Wu. It is you guys that make my Groningen life much happier.

I gratefully acknowledge financial support from China Scholarship Council (CSC).

Finally, I reserve my deepest gratitude to my family. Thank my sister and my brother for their support. Thank my parents and my wife for their unlimited love.

Chapter 1

Introduction

This thesis explores planning from a logical point of view. Planning theory is a central area in Artificial Intelligence (AI), and it is concerned with finding a plan (such as a sequence of actions) to achieve goals. In AI, this theory consists of several areas each having its own restrictive assumptions about the kind of setting in which the planning takes place. In this thesis, we are mainly concerned with the area of planning under uncertainty. Here the agent is uncertain about his situation. There are two main topics of this thesis. One is to build a logical framework for capturing how the agent's uncertainty evolves in these planning problems. The other topic, since these planning problems bring us some new idea of what it means to *know how* to achieve a goal, is to model the new understanding of knowing how in logical systems.

1.1 Planning under uncertainty

Conformant planning is the simplest version of planning under uncertainty. Before we introduce it, we first explain the notion of a transition system. A (labelled) *transition system* is a tuple $\langle S, Act, R\rangle$ where S is a set of states, Act is a set of actions (or labels), and R is a set of labelled transitions (i.e. a subset of $S \times Act \times S$). If S is finite, it is called a finite transition system. The fact that $(s, a, t) \in R$ (also written as $s \xrightarrow{a} t$, or $t \in R_a(s)$) represents that there is a transition from s to t with label a. Intuitively, it means that performing the action a in state s might result in state t. A transition system is *deterministic* if the next state is completely determined by the current state and the action executed by the agent, which is formally expressed as that $(s, a, t), (s, a, t') \in R$ implies $t = t'$ (so that R is a partial function). Otherwise, it is *nondeterministic*.

Conformant planning Given a nondeterministic transition system $\langle S, Act, R\rangle$, an initial uncertainty set $U \subseteq S$, and a goal set $G \subseteq S$, *conformant planning* is to find a plan which is a finite linear sequence of actions such that performing the plan in each state of U will never fail and always result in states in G (cf. Ghallab et al. (2004)). Such an action sequence is called a *conformant plan*.

In conformant planning, it is assumed that the agent has no sensors, i.e. the environment is unobservable. There are two kinds of uncertainty in conformant planning. The location of the agent can be uncertain since the environment is unobservable, and the results of performing an action in a state can be uncertain since the transition system is nondeterministic. If we restrict the transition system to be deterministic and the environment is fully observable (which leads the initial uncertainty set to be a singleton), we end up with a problem of *classical planning*.

Here is an example of conformant planning. Figure 1.1 depicts a transition system which can be seen as a map, where s_i are places connected by corridors (r) or stairs (u).[1] Let the initial uncertainty set be $\{s_2, s_3\}$, that is, the agent is uncertain whether he is now in s_2 or s_3, and let the goal set be $\{s_7, s_8, s_5\}$. From Figure 1.1, we can see that performing ru in s_2 will lead to s_7, and that performing ru in s_3 will lead to s_8. Both s_7 and s_8 are in the goal set, so the planning problem is solved. Please note that the environment is unobservable in conformant planning, which means there is no feedback during the execution of plans. In this example, if the agent is initially uncertain about s_2 or s_3, he is uncertain about s_3 or s_4 after moving right (r).

$$\begin{tikzcd}
 & s_6 & s_7 & s_8 & \\
s_1 \arrow[r, "r"] & s_2 \arrow[r, "r"] \arrow[u, "u"] \arrow[ur, "u"] & s_3 \arrow[r, "r"] \arrow[u, "u"] & s_4 \arrow[r, "r"] \arrow[u, "u"] & s_5
\end{tikzcd}$$

Figure 1.1

As mentioned previously, conformant planning is the simplest version of planning under uncertainty. If we extend the transition system of conformant planning with probabilities, we are in the area of *conformant probabilistic planning*.

Conformant probabilistic planning A probabilistic transition system is a quadruple $\langle S, Act, R, Pr\rangle$, where $\langle S, Act, R\rangle$ is a finite transition system and $Pr : R \to (0, 1]$ is a transition probability function. Given a probabilistic transition system, an initial uncertainty set U, an initial belief state $I : U \to (0, 1]$ which is a probability distribution over U, and a goal set G, *conformant probabilistic planning* is to find a plan which is a finite sequence of actions such that the probability of reaching the goal by performing the plan is no less than a given threshold $\delta \leq 1$. If the threshold δ is 1, we end up with a problem of conformant planning.

The following is a conformant probabilistic planning problem. Figure 1.2 depicts a probabilistic transition system and an initial belief state of which the domain is $U = \{s_1, s_2\}$. Let the goal set be $\{s_3, s_5\}$. There is no conformant plan in this example. We can see that the probability of reaching the goal after performing a is $0.5 \times 0.8 = 0.4$, and that the probability of reaching the goal after performing b is $0.5 \times 1 = 0.5$. Therefore, if the threshold is 0.5, the action b is a solution.

Just as conformant planning, there is no observability in conformant probabilistic planning. If we extend conformant planning with partial observability, we are in the area of *contingent planning*.

Contingent planning A partially observable model is a tuple $\langle S, Act, R, O\rangle$, where $\langle S, Act, R\rangle$ is a transition system and O is an equivalence relation on S which reflects

[1] This is a variant of the running example used in Wang and Li (2012).

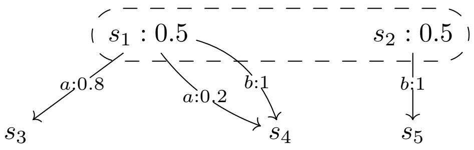


Figure 1.2

the features of the environment that can be observed [2]. Given a partially observable model, an initial uncertainty set U, and a goal set G, a contingent plan is a partial function from $\mathcal{P}(S)$ to Act, which tells the agent which action to execute based on his current uncertainty set, such that if the agent acts accordingly then he is guaranteed to reach the goal after finitely many steps.

Partial observability is more general than full observability and non-observability because if O is the identity relation then it is fully observable and if O is the universal relation $S \times S$ then it is unobservable. Please note that a contingent plan only makes sense under partial observability. In an unobservable model (i.e. $O = S \times S$), there must be a conformant plan if there exists a contingent plan.

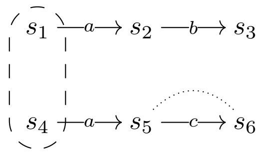


Figure 1.3

Figure 1.3 depicts a partially observable model and an initial uncertainty set $U = \{s_1, s_4\}$, where the equivalence relation O is represented by the dotted line and we omit the reflexive dotted arrows. Please note that the agent is certain about his location if he performs a and, for example arrives at s_5, at which time his uncertainty set will update to be $\{s_5\}$. Even though he would be uncertain about s_5 or s_2 without observation after a, he will distinguish these two states by observation.

Given a partially observable model and an initial uncertainty set, there is an extended model in which all the possible uncertainty sets are explicit. For example, the extended model of Figure 1.3 is depicted in Figure 1.4. Let the goal set be $G = \{s_3, s_6\}$. We can see that the partial function $f = \{\{s_1, s_4\} \mapsto a, \{s_2\} \mapsto b, \{s_5\} \mapsto c\}$ is a contingent plan.

We have introduced three kinds of planning problems under uncertainty: conformant planning, conformant probabilistic planning, and contingent planning. Both conformant probabilistic planning and contingent planning are extensions of conformant planning. Besides these, there are many other versions of planning in AI (cf. Ghallab et al. (2004)), such as planning that requires a successful solution to be finished in restricted time. In this thesis, we are mainly concerned with these three kinds of planning because an agent's uncertainty plays an important role in these planning problems.

[2]In the standard formulation of contingent planning, O is an observation function from S to $\mathcal{P}(T)$ where T is a set of tokens, but we can always generate an equivalence relation based on the observation function.

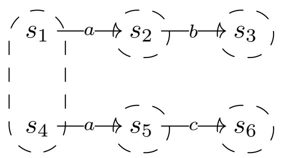


Figure 1.4

1.2 Logics for planning

Dynamic epistemic logic (DEL) (cf. e.g., van Ditmarsch et al. (2007)) is a standard framework for reasoning about knowledge and change. In DEL, there is an epistemic model which describes the initial knowledge situation, and an operation called *product update* which computes a new epistemic model based on the old epistemic model and an event model. The product update describes the change of knowledge. In recent years, there has been a growing interest in applying DEL in planning under uncertainty, using the update of knowledge to capture the change of uncertainty during the execution of plans. In these DEL-based planning frameworks, states in the uncertainty set constitute an epistemic model, actions are event models, and the state transitions are implicitly encoded by the update product (cf. Bolander and Andersen (2011); Löwe et al. (2011)).

One advantage of this DEL-based approach is its expressiveness in handling multi-agent planning with knowledge goals (cf. e.g., Bolander and Andersen (2011); Löwe et al. (2011); Andersen et al. (2012); Aucher (2012); Yu et al. (2013); Pardo and Sadrzadeh (2013); Jensen (2014); Bolander et al. (2015); Muise et al. (2015)), while the traditional AI planning focuses on the single-agent case. In particular, the event models of DEL (cf. Baltag and Moss (2004)) are used to handle non-public actions that may cause different knowledge updates for different agents. However, this expressiveness comes at a price, as shown in Bolander and Andersen (2011); Aucher and Bolander (2013), that multi-agent epistemic planning is undecidable in general. Many interesting decidable fragments are found in the literature (Bolander and Andersen (2011); Löwe et al. (2011); Yu et al. (2013); Andersen et al. (2015)), which suggests that single-agent cases and some restrictions on the structure of event models are the keys to decidability.

Another logic framework for planning under uncertainty is the epistemic propositional dynamic logic (EPDL) proposed in Yu et al. (2016) co-authored with Qu Yu and Yanjing Wang. The model of EPDL is simply a transition system with initial uncertainty as in the example shown in Figure 1.1. Even though there is no event model in EPDL, EPDL still adopts the idea of DEL by interpreting actions in the semantics as an update on the uncertainty of the agent. EPDL also follows the idea of *planning as model checking* explored in Giunchiglia and Traverso (2000); van der Hoek and Wooldridge (2002); Jamroga and Ågotnes (2007). The language of EPDL is very powerful, and it reduces the problem whether a planning problem has a solution to a model checking problem whether a EPDL formula is true in a model.

One impressive feature of the EPDL approach is that standard conformant planning is generalized in the EPDL framework but the generalized conformant planning is equally hard as the standard conformant planning over explicit transition systems. EPDL generalizes the standard conformant planning problem in AI (over transition systems) in two ways, w.r.t. the goal and also the constraint on the desired plan. First, the planning goal can be any EPDL formula. Since in EPDL regular operators are used to construct complex actions, we can express the goal "the agent knows that φ is forever true" by $\mathcal{K}[(a+b)^*]\varphi$ (assuming there are only two actions a and b). Second, procedural constraints on the desired plan specified by regular expressions can be imposed. For example, we can require that only action sequences where b cannot be executed before a are desired, which is expressed by $(a^*; b^*)$. Furthermore, as pointed in Li et al. (2017), EPDL can be naturally extended to cover contingent planning. EPDL is not included in this thesis. For further reading about EPDL, please see Wang and Li (2012); Yu et al. (2016); Li (2015, 2016).

Both approaches above (DEL and EPDL) are concerned with conformant planning and contingent planning. In Chapter 2 of this thesis, we propose a logic framework to deal with the change of belief state in conformant probabilistic planning. We adopt the idea of probabilistic dynamic epistemic logic (PDEL). PDEL, presented in Appendix A.2, is a combination of probability logic and public announcement logic, which models the change of belief due to announcements. Similar to EPDL, the logic proposed in Chapter 2 is over a probabilistic transition system and an initial belief state, where actions are interpreted in the semantics as an update on the belief state.

1.3 Logics inspired by planning

Epistemic logic presented in Appendix A.1 is a logical formalism of propositional knowledge, which is the knowledge expressed by the phrase of knowing that p. Epistemic logic is commonly accepted and widely applied in many fields, such as game theory, theoretical computer science, artificial intelligence, and so on. Besides propositional knowledge, procedural knowledge is another kind of knowledge often discussed in epistemology, which is the knowledge expressed by the phrase of *knowing how* to do something. However, there is no common consensus on how to formalize knowing how (cf. the recent surveys Gochet (2013) and Ågotnes et al. (2015)).

Dating back to McCarthy and Hayes (1969); McCarthy (1979); Moore (1985); Singh (1994); Lespérance et al. (2000); van der Hoek et al. (2000), researchers have been looking at knowing how in the setting of propositional knowledge and actions, but the difficulty is that simply combining the existing logics for knowing that and ability does not lead to a genuine notion of knowing how (cf. Jamroga and Ågotnes (2007); Herzig (2015)). For example, knowing how to achieve p is not equivalent to knowing that there exists a strategy to make sure that p. Let $\varphi(x)$ express that x is a way to make sure some goal is achieved. There is a crucial distinction between the *de dicto* reading of knowing how ($\mathcal{K}\exists x\varphi(x)$) and the desired *de re* reading ($\exists x\mathcal{K}\varphi(x)$) (cf. Stanley and Williamson (2001); Jamroga and van der Hoek (2004); Ågotnes (2006)). The latter implies the former, but not the other way round. 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 seeing-to-it-that (STIT) logic (Broersen and Herzig (2015)).

Inspired by the idea of conformant planning, Wang (2015a, 2016) proposes a new logic of knowing how presented in Appendix A.3, in which agents' knowing how to achieve a goal is interpreted as having a conformant plan for the goal. This approach is in line with the *de re* reading of knowing how, but he introduces a single new modality $\mathcal{K}h$ of (goal-directed) knowing how, instead of breaking it down into other modalities. The modality $\mathcal{K}h$ is a binary modality. The model of this logic is a transition system extended with an assignment, and the modality $\mathcal{K}h$ is universal on the semantics. For example, the formula $\mathcal{K}h(p, q)$, which reads as agents' knowing how to achieve q-states from p-states, is true in the model depicted in Figure 1.5, if and only if there is a conformant plan with given initial uncertainty set $U = \{s_2, s_3\}$ and goal set $G = \{s_7, s_8, s_5\}$.

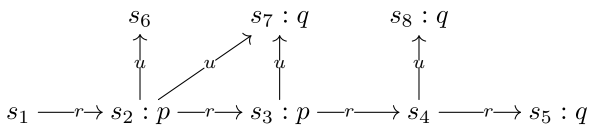


Figure 1.5

In Chapter 3 and Chapter 4 of this thesis, we extend the work of Wang (2015a, 2016). In real-life contexts, constraints on how we achieve the goal often matter. For example, we want to know how to win the game by playing fairly; people want to know how to be rich without breaking the law. In Chapter 3, we generalize the modality $\mathcal{K}h$ to be a ternary knowing how operator to express that the agent knows how to achieve φ given ψ while maintaining χ in-between. What is more, we think that, in our everyday life, the requirement of a conformant plan might be too strong for knowing-how. For example, considering Figure 1.5, let the initial uncertainty set remain the same as in Figure 1.1, i.e. $U = \{s_2, s_3\}$, but the goal is $\{s_5\}$. Intuitively we still say we know how to achieve s_5 because we can get there by moving right (r) at most three times. However, the plan rrr is not a conformant plan since performing it in s_3 will fail. We call such a plan "weak conformant plan". In Chapter 4, we weaken the interpretation of the formula $\mathcal{K}h(p, q)$ as having a *weak conformant plan* for achieving q given p.

Inspired by the idea of contingent planning, Chapter 5 extends the standard epistemic logic with a unary knowing-how modality $\mathcal{K}h^{\mathsf{s}}$, in which the formula $\mathcal{K}h^{\mathsf{s}}\varphi$ is interpreted as having a contingent plan to achieve φ. Since a contingent plan is also called a strategy, we call the logic *strategically knowing-how logic* (SKH). The model of SKH is a transition system extended with an equivalence relation, like Figure 1.4. Chapter 6 is a preparation for making SKH epistemically dynamic. As we know, extending epistemic logic with public announcement operators $[\varphi]$ is a natural way to make the logic dynamic and to reason about knowledge change. However, it does not work for the logic SKH because updating a model of SKH with $[\varphi]$ might remove some states from the model, which will change the basic transition system of the model. Besides public announcement logic, arrow update logic presented in Appendix A.4 proposes another way to reasoning about knowledge change by removing epistemic accessibilities in a model. This method can be applied to SKH. The limitation is that the standard arrow update logic can only deal with public announcements. To reason about knowledge change due to private announcements, Chapter 6 extends the standard arrow update logic.

1.4 Overview

Each main chapter is based on a published or submitted paper. To keep each chapter independently readable, there is some overlap among chapters. The contribution of each chapter is briefly summarized as follows.

Chapter 2, *a logic for conformant probabilistic planning*, is an extended version of joint work with Barteld Kooi and Yanjing Wang. In this chapter, we introduce a logic framework that can be applied to conformant probabilistic planning. Conformant probabilistic planning is to find a linear plan (a sequence of probabilistic actions) such that the probability of achieving the goal (a certain condition) after executing the plan is no less than a given threshold probability δ. Our logical framework can trace the change of the belief state of the agent (which is a probability distribution over states) during the execution of the plan. With this logic, we can enrich the standard conformant probabilistic planning by formulating the goal as a dynamic epistemic logic formula. We provide a complete axiomatization of the logic. Moreover, this chapter shows that the logic is decidable.

Chapter 3, *knowing how with intermediate constraints*, is an extended version of joint work with Yanjing Wang (Li and Wang (2017)). In this chapter, we propose a triple knowing-how operator to express that the agent knows how to achieve φ given ψ while maintaining χ in-between. It generalizes the knowing-how logic which is presented in Appendix A.3. We give a sound and complete axiomatization of this logic. What is more, this chapter introduces a filtration method on the canonical model. By the filtration method, it is shown that this logic has a small model property, and thus the logic is decidable.

Chapter 4, *knowing how with weak conformant plans*, is an updated version of the paper Li (2017). This chapter proposes a weaker but more realistic semantics to the modality of the knowing-how logic presented in Appendix A.3. According to this semantics, an agent knows how to achieve φ given ψ if (s)he has a finite linear plan by which (s)he can always end up with a φ-state when the execution of the plan terminates, whether or not each part of the execution has been completed. This weaker interpretation of the knowing-how modality results in a weaker logic than the knowing-how logic. The composition axiom of the knowing-how logic is no longer valid. This chapter also presents a sound and complete axiomatic system and proves that this logic is decidable.

Chapter 5, *strategically knowing how*, is an elaborate version of joint work with Raul Fervari, Andreas Herzig, and Yanjing Wang (Fervari et al. (2017)). In this chapter, we extend the standard epistemic logic of knowing that with a new knowing-how operator. The semantics of the new knowing-how 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 that 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 both knowledge-that and knowledge-how.

Chapter 6, *privacy in arrow update logic*, is based on a short paper presented on the conference of *Advances in Modal Logic 2014*. This chapter develops the arrow update logic. Arrow update logic presented in Appendix A.4 is a theory of epistemic access elimination that can be used to reason about information change. In the arrow update logic, it is common knowledge among agents how each will process incoming information. This chapter develops the basic theory of arrow update logic to deal with private and semi-private announcements. In this framework, the information is private for an agent group. This chapter also proposes a labelled tableau calculus for this logic and shows that this logic is decidable.

Chapter 2

A logic for conformant probabilistic planning[1]

2.1 Introduction

Automated planning is a branch of artificial intelligence concerned with devising a plan, which might be a strategy or an action sequence, to achieve some goals. Automated planning technology is widely applied in a variety of areas, ranging from controlling the operations of spacecraft to playing the game of bridge. Classical planning, which is the simplest form of automated planning, is the problem of finding a linear action sequence in a deterministic transition system such that executing the plan in the initial state will achieve the goal (cf. Ghallab et al. (2004)). There are two important simplifying assumptions for classical planning: *determinacy* and *full observability*. Full observability indicates that agent has complete knowledge about the system and the state in which the system starts, which means that the set of initial states from which the plan starts is a singleton.

Conformant planning generalizes classical planning by relaxing these two restrictions, namely that it allows lack of knowledge of position in the system (and lack of ability to observe where agent is located) and it allows actions to be non-deterministic. The former means that the set of initial states needs no longer necessarily be a singleton (and corresponds to the agent's initial uncertainty) but also means that one cannot attain certainty regarding one's whereabouts based on observation during plan execution. A conformant plan is an action sequence to guarantee the agent's arrival at one of the goal states no matter what initial state the plan starts from and no matter how the (non-deterministic) plan is executed (cf. Ghallab et al. (2004)). Since conformant planning brings out an agent's uncertainty, rather than from the traditional AI approaches, we can profit from an epistemic-logical perspective on conformant planning.

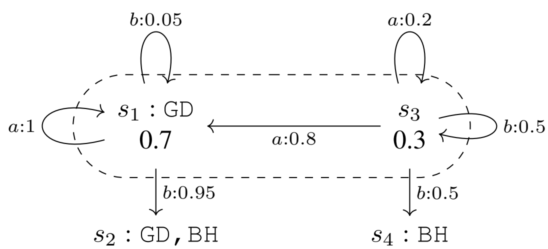
Epistemic planning is concerned with planning under uncertainty. Dynamic epistemic logic (DEL) (cf. van Ditmarsch et al. (2007)) can be used as a formal framework to deal with epistemic planning and can provide a useful generalization of conformant planning by allowing the planner to formulate knowledge goals within a formal lan-

[1] This is an extended version of joint work with Barteld Kooi and Yanjing Wang.

guage. Moreover, by applying DEL in planning, we can handle multi-agent planning with knowledge goals (cf. e.g., Andersen et al. (2012); Aucher (2012); Bolander and Andersen (2011); Bolander et al. (2015); Jensen (2014); Löwe et al. (2011); Muise et al. (2015); Pardo and Sadrzadeh (2013); Yu et al. (2013)). In particular, the event models of DEL (cf. Baltag and Moss (2004)) are used to handle non-public actions that may yield different knowledge updates for different agents. Within such DEL-based approaches, Wang and Li (2012) proposed a dynamic epistemic logic over transition systems with initial uncertainty. This framework is extended in Yu et al. (2016) with propositional dynamic logic programs in such a way that the standard conformant planning problem has the form of a model checking problem in this framework.

Conformant probabilistic planning (CPP) is another significant generalization of conformant planning. The demands that conformant planning puts on the solution may be too strong, in the sense that a solution has to be found for sure. It may be the case that a solution in this sense is impossible while there may still be some plan that leads to a goal state with very high probability. Conformant probabilistic planning generalizes conformant planning by relaxing the conditions on solutions to planning problems. The framework of CPP is enriched with a probability distribution over states and tells us the probability that a certain action will lead to a certain successor state. In this way probabilistic goals can be set that are easier to achieve than non-probabilistic goals, but will still be such that they do provide a satisfactory plan that will lead to a goal state with acceptable probability.

Let us consider the following toy example of a planning problem where we need to take probability into account.[2] Take a robot whose gripper is possibly wet. The gripper needs to hold a block, but gripping a block while the gripper is slippery is more difficult than when it's dry. This can be modeled in a transition system with probabilities:



There are two propositions: GD stands for gripper-dry and BH for block-held, and two actions: a stands for drying and b for picking up. We model the initial belief state by a probability distribution over the states of the system B, which assigns the following probabilities: $B(s_1) = 0.7$ and $B(s_3) = 0.3$. The action a dries a dry gripper with probability 1, but make a wet gripper dry with probability 0.8. The action b picks up the block with probability 0.95 if the gripper is dry and with probability 0.5 if the gripper is wet. It is impossible to find a plan in this example that will guarantee that after executing the plan the robot will hold a block. But for practical purposes it may be enough to find a plan to hold the block, which succeeds at least 90% of the time.

Conformant probabilistic planning is well studied in AI literature (cf. e.g., Kushmerick et al. (1995); Hyafil and Bacchus (2003, 2004); Bryce et al. (2008); Taig and

[2]This is a variant of the Slippery Gripper example reported in Kushmerick et al. (1995); Hyafil and Bacchus (2003).

Brafman (2013)). A variety of algorithms is developed to solve CPP problems, and each algorithm has its own advantage given certain assumptions about the system. Generally, the probability of achieving a goal state t by executing a plan $\pi = a_1 \cdots a_n$ is calculated in the following way (cf. Hyafil and Bacchus (2004)):

$$\mu_\pi(t) = \sum_{\{s_0 \cdots s_n \mid \forall 1 \leq i \leq n : s_{i-1} \xrightarrow{a_i} s_i, s_n = t\}} B(s_0) \times Pr(s_0, a_1, s_1) \times \cdots \times Pr(s_{n-1}, a_n, s_n)$$

where $Pr(s_{i-1}, a_i, s_i)$ means the probability of reaching s_i after executing a_i at s_{i-1}. For example, in the Slippery Gripper example, the probability of achieving s_2 by executing ab is the following:

$$\begin{aligned}
&\mu_{ab}(s_2)\\
&= B(s_1) \times Pr(s_1, a, s_1) \times Pr(s_1, b, s_2) + B(s_3) \times Pr(s_3, a, s_1) \times Pr(s_1, b, s_2)\\
&= 0.7 \times 1 \times 0.95 + 0.3 \times 0.8 \times 0.95\\
&= 0.665 + 0.228 = 0.893
\end{aligned}$$

To achieve a higher probability of holding the block, the robot has to (try to) dry the gripper twice.

In this chapter, we build a logic framework over the CPP models so that we can enrich the standard CPP problems with much more powerful goal formulas and investigate the reasoning about actions and plans in CPP. Roughly speaking, we generalize single-agent epistemic planning and conformant probabilistic planning into one framework. Our approach combines conformant probabilistic planning and probabilistic dynamic epistemic logic (PDEL), which is presented in Appendix A.2 (cf. Kooi (2003); van Benthem (2003); van Benthem et al. (2009)). In our language, there are two kinds of modalities: action modalities $[a]$ and probability modalities B_π. As in PDEL, the action modality is interpreted in the semantics as an update on the belief state which is a probability distribution. The probability modality B_π is interpreted as μ_π like in the CPP literature.

Our framework can also distinguish different conformant probabilistic plans more precisely. Consider the model $\mathcal{M}$ depicted in Figure 2.1. Let the goal be to reach p-

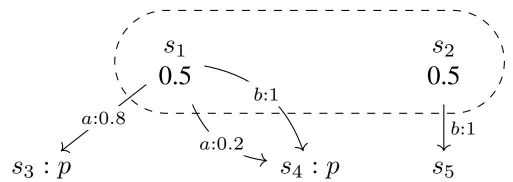


Figure 2.1: $\mathcal{M}$

states with threshold probability 0.5, then both a and b are solutions. However, although $\mu_a(p) = \mu_b(p) = 0.5$, intuitively we feel that there are differences between these two solutions. Plan a cannot always be successfully executed, but the agent is guaranteed to reach a p-state when a is successfully executed. Contrary to the plan a, the plan b can always be successfully executed, but there is only 50% probability to achieve p-states

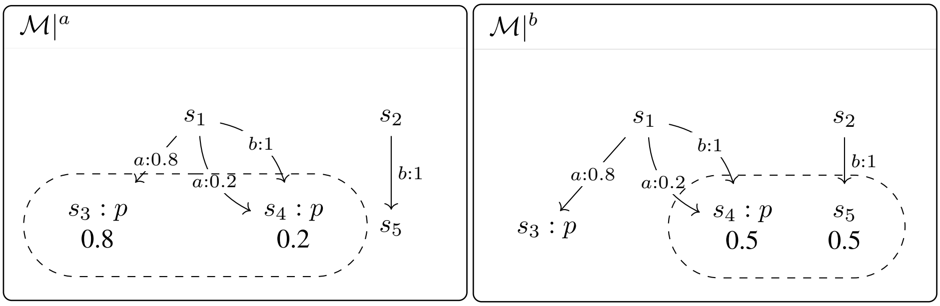


Figure 2.2

by doing b. In our framework, this can be expressed as $B_a\top = 0.5 \wedge [a](B_\epsilon p = 1)$ and $B_b\top = 1 \wedge [b](B_\epsilon p = 0.5)$. Therefore, if we care more about the probability of achieving goals given the successful execution of the plan, a is a better solution than b.

Our framework is significantly different from PDEL since PDEL works with event models rather than transition systems. Moreover, there is no syntax that is directly linked to μ_π. This is because μ_π can be seen as a conditional probability weighed by the probability of π being successfully executed (and in that sense it is closer to a prior probability), namely:

$$\mu_\pi(t) = Pr(t \mid ex(\pi)) \cdot Pr(ex(\pi))$$

where $Pr(ex(\pi))$ is the probability that π can be successfully executed. $Pr(t \mid ex(\pi))$, which is calculated in the updated model given the execution of π, can be expressed in PDEL, but $Pr(ex(\pi))$ cannot be expressed in PDEL since PDEL is only concerned with probabilities expressing belief and not with the probability of actions. However, $Pr(ex(\pi))$ is a combination of belief probability and action probability.

The rest of this chapter is organized as follows: Section 2.2 introduces the language and semantics, and also defines conformant probabilistic planning in terms of our logic framework; Section 2.3 presents the axiomatics of this logic; Section 2.4 proves its completeness; the last section concludes with some future directions.

2.2 The logic LCPP

In this section we introduce the logic of conformant probabilistic planning, and we denote the logic as LCPP. Besides a (somewhat limited) dynamic logic of action, the language of LCPP also has linear inequalities of weighted probabilistic terms, which express the probability that a sequence of actions reaches a certain set of states. This language differs from the usual languages of PDEL in the sense that here actions are atomic and do not have an internal structure which explicates how it changes information. Also probabilistic expressions are indexed by sequences of actions, rather than having no index. Of course in PDEL probability terms are indexed by agents. To keep things simple in this chapter we focus on the single-agent case.

Definition 2.2.1 (Language) *Let a countable set of propositional variables* $\mathbf{P}$ *and a* *finite set of actions* $\mathbf{Act}$ *be given. The language* $\mathcal{L}_{LCPP}$ *is defined as the following BNF:*

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid [a]\varphi \mid q_1 B_{\pi_1}\varphi_1 + \cdots + q_n B_{\pi_n}\varphi_n \geq q$$

where $p \in \mathbf{P}$, $a \in \mathbf{Act}$, $\pi_i \in \mathbf{Act}^*$, *i.e. a finite string (possible empty) of actions and* $q, q_i \in \mathbb{Q}$ *for each* $1 \leq i \leq n$.

Besides the usual abbreviations, we have the following.

$$\begin{array}{lcl}
\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \geq q & := & q_1 B_{\pi_1}\varphi_1 + \cdots + q_n B_{\pi_n}\varphi_n \geq q \\
q_1 B_{\pi_1}\varphi_1 \geq q_2 B_{\pi_2}\varphi_2 & := & q_1 B_{\pi_1}\varphi_1 + (-q_2) B_{\pi_2}\varphi_2 \geq 0 \\
\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \leq q & := & \sum_{i=1}^n (-q_i) B_{\pi_i}\varphi_i \geq (-q) \\
\sum_{i=1}^n q_i B_{\pi_i}\varphi_i < q & := & \neg(\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \geq q) \\
\sum_{i=1}^n q_i B_{\pi_i}\varphi_i > q & := & \neg(\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \leq q) \\
\sum_{i=1}^n q_i B_{\pi_i}\varphi_i = q & := & (\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \geq q) \wedge (\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \leq q) \\
\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \neq q & := & (\sum_{i=1}^n q_i B_{\pi_i}\varphi_i > q) \vee (\sum_{i=1}^n q_i B_{\pi_i}\varphi_i < q) \\
B_\pi\varphi = B_{\pi'}\varphi' & := & 1B_\pi\varphi - 1B_{\pi'}\varphi' = 0 \\
K\varphi & := & B_\epsilon\varphi = 1 \\
\hat{K}\varphi & := & \neg K\neg\varphi \\
\langle a\rangle\varphi & := & \neg[a]\neg\varphi \\
\langle a_1 \cdots a_n\rangle\varphi & := & \langle a_1\rangle \cdots \langle a_n\rangle\varphi \\
(\!|a|\!)\varphi & := & [a]\varphi \wedge \langle a\rangle\varphi \\
(\!|a_1 \cdots a_n|\!)\varphi & := & (\!|a_1|\!) \cdots (\!|a_n|\!)\varphi \\
[a_1 \cdots a_n]\varphi & := & [a_1] \cdots [a_n]\varphi
\end{array}$$

We call formula of the form $\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \bowtie q$ *as* probability formula*, where* $\bowtie$ *is one of* $\leq, \geq, <, >, =, \neq$.

Let us explain how to read the formulas of the language. Propositional variables such as p express basic properties of a world, such as “the coin landed heads”. Then we have standard negation and conjunction. We read formulas of the form $[a]\varphi$ as “after all executions of action a it is the case that φ”. In order to read linear inequality formulas of the form $q_1 B_{\pi_1}\varphi_1 + \cdots + q_n B_{\pi_n}\varphi_n \geq q$ we first explain how to read $B_\pi\varphi$. The essential idea is that it represents the probability of getting φ using π. More precisely, it consists of two parts: the probability of φ given the successful execution of π, and the probability of the successful execution of π. Roughly speaking, we have $B_\pi\varphi = Pr(\varphi \mid ex(\pi)) \cdot Pr(ex(\pi))$. As will become more clear after introducing the semantics, $Pr(\varphi \mid ex(\pi))$ will be calculated in the updated model given the execution of π. In other words, $B_\pi\varphi$ is the non-normalized probability of φ in the model you get by executing π. If π is not executable, this should be zero, and if it is executable it is the probability of φ in the updated model multiplied by the probability of the executability of π in order to undo the normalization. Now the linear equality is where we take a sum of these terms multiplied by rational numbers and compare this sum to a rational number. The semantics of Definition 2.2.9 will make this precise.

The language is interpreted on models which are in a sense probabilistic transition systems. There are two kinds of probabilistic elements in these models. There is a prior probability distribution representing the initial uncertainty of the agent, and there is a probability function which indicates for each state and each action that can be executed at that state which probability one has of reaching some other state. There is one additional difference between these models and the models that also appear in some of the CPP literature, namely that we allow actions to be unexecutable, and we allow the initial uncertainty only to be concerned with a strict subset of the set of all states.

Definition 2.2.2 (Model) *A model $\mathcal{M}$ is a tuple $\langle S^{\mathcal{M}}, R^{\mathcal{M}}, Pr^{\mathcal{M}}, I^{\mathcal{M}}, B^{\mathcal{M}}, V^{\mathcal{M}}\rangle$ such that*

- *$S^{\mathcal{M}} \neq \emptyset$, a (finite) set of states;*[3]
- *$R^{\mathcal{M}} \subseteq S^{\mathcal{M}} \times \mathbf{Act} \times S^{\mathcal{M}}$, a non-deterministic execution relation for each action;*
- *$Pr : R^{\mathcal{M}} \to \mathbb{Q}^+$ is a probability function that expresses the probability that an action will lead to another state, such that for all $a \in \mathbf{Act}$ it holds that $\sum_{t \in R_a^{\mathcal{M}}(s)} Pr^{\mathcal{M}}(s, a, t) = 1$;*
- *$I^{\mathcal{M}}$ is a non-empty subset of $S^{\mathcal{M}}$, consisting of those states that the agent considers possible;*
- *$B^{\mathcal{M}} : I^{\mathcal{M}} \to \mathbb{Q}^+$ is a probability distribution over states in $I^{\mathcal{M}}$ expressing the probability of being the true initial state such that $\sum_{s' \in I^{\mathcal{M}}} B^{\mathcal{M}}(s') = 1$;*
- *$V^{\mathcal{M}} : \mathbf{P} \to \mathcal{P}(S^{\mathcal{M}})$, a valuation function indicating for each propositional variable in which set of worlds it holds.*

For each $s \in I^{\mathcal{M}}$, $(\mathcal{M}, s)$ is a pointed model.

Given $\mathcal{M}$, $(s, a, t) \in R^{\mathcal{M}}$ is also denoted as $s \xrightarrow{a} t$, $(s, t) \in R_a^{\mathcal{M}}$ or $t \in R_a^{\mathcal{M}}(s)$.

Before we provide the semantics, we first provide the notions needed to define how models are updated by executing a sequence of actions, since we need those models to interpret actions and probabilistic statements. First, we define the semantic structure that is associated with a sequence of actions, called the set of execution paths.

Definition 2.2.3 *Given $\mathcal{M}$, $\pi = a_1 \cdots a_n$, we call $s_0 a_1 \cdots s_n$, which is an alternating sequence of states and actions, an* execution path *of π in $\mathcal{M}$ if $s_0 \in I^{\mathcal{M}}$ and $s_{i-1} \xrightarrow{a_i} s_i$ for each $1 \leq i \leq n$. If the action sequence is obvious, we also write the execution path $s_0 a_1 \cdots s_n$ as a sequence of states $s_0 \cdots s_n$. The set of execution paths of π in $\mathcal{M}$ is denoted as $EP_{\mathcal{M}}(\pi)$.*

After executing a sequence π, the probability the agent assigns to the states of the model changes. Let $I^{\mathcal{M}}|^a$ be the set $\{t \in S^{\mathcal{M}} \mid s \xrightarrow{a} t \text{ for some } s \in I^{\mathcal{M}}\}$, and $I^{\mathcal{M}}|^{\pi} = I^{\mathcal{M}}|^{a_1} \cdots |^{a_n}$ where $\pi = a_1 \cdots a_n$. We use the following auxiliary notion to update this probability.

Definition 2.2.4 *Given $\mathcal{M}$ and $\pi = a_1 \cdots a_n \in \mathbf{Act}^*$, the function $\mu_{\pi}^{\mathcal{M}} : I^{\mathcal{M}}|^{\pi} \to (0, 1]$ is defined as follows: for each $t \in I^{\mathcal{M}}|^{\pi}$,*

$$\mu_{\pi}^{\mathcal{M}}(t) = \sum_{\{s_0 \cdots s_n \in EP_{\mathcal{M}}(\pi) \mid s_n = t\}} (B^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}}(s_{i-1}, a_i, s_i))$$

Given $T \subseteq I^{\mathcal{M}}|^{\pi}$ and π, let $\mu_{\pi}^{\mathcal{M}}(T) = \sum_{t \in T} \mu_{\pi}^{\mathcal{M}}(t)$, especially, $\mu_{\pi}^{\mathcal{M}}(\emptyset) = 0$.

[3]The restriction to a finite set of states is to make the presentation more simple. We could easily remove this restriction and use sigma-algebras and fully general probability theory, but this would only distract from the issues we are exploring in this chapter.

Intuitively, $\mu_\pi^{\mathcal{M}}(t)$ stands for the agent's belief degree of reaching t after the agent performs π.

Remark 2.2.5 *Similar to the forward algorithm for computing the probability of a particular observable sequence in Hidden Markov Models (cf. e.g., Rabiner (1990)), we can also compute $\mu_\pi^{\mathcal{M}}(t)$ recursively by computing $\mu_{\pi'}^{\mathcal{M}}(t')$ for all the initial segments π' of π and the relevant states t'.*

The updated probability of the agent applies to a possibly updated set of states that the agent considers possible. Now we define the updated probability of the agent.

Definition 2.2.6 *Given $\mathcal{M}$ and $\pi = a_1 \cdots a_n \in \mathbf{Act}^*$ such that $I^{\mathcal{M}}|^\pi \neq \emptyset$, function $B^{\mathcal{M}}|^\pi : I^{\mathcal{M}}|^\pi \to \mathbb{Q}$ is defined as follows: for each $t \in I^{\mathcal{M}}|^\pi$,*

$$B^{\mathcal{M}}|^\pi(t) = \frac{\mu_\pi^{\mathcal{M}}(t)}{\mu_\pi^{\mathcal{M}}(I^{\mathcal{M}}|^\pi)}$$

Note that in this definition both the numerator and the denominator are non-zero given the way we set things up. Note that by assuming that $I^{\mathcal{M}}|^\pi$ is non-zero it follows that $EP_{\mathcal{M}}(\pi)$ is non-empty. Since we assumed that the probability functions in the model only assign positive probabilities and that t is in $I^{\mathcal{M}}|^\pi$, both numerator and denominator are non-zero. More formally:

Proposition 2.2.7 *Given $\mathcal{M}$, $\pi = a_1 \cdots a_n$ and $I|^\pi \neq \emptyset$, we have that $B^{\mathcal{M}}|^\pi$ is a probability function from $I^{\mathcal{M}}|^\pi$ to $\mathbb{Q}^+$ and $\sum_{t \in I^{\mathcal{M}}|^\pi} B^{\mathcal{M}}|^\pi(t) = 1$.*

Given all these definitions, it is now easy to define the updated model.

Definition 2.2.8 *Given model $\mathcal{M} = \langle S^{\mathcal{M}}, R^{\mathcal{M}}, Pr^{\mathcal{M}}, I^{\mathcal{M}}, B^{\mathcal{M}}, V^{\mathcal{M}} \rangle$ and $I^{\mathcal{M}}|^\pi \neq \emptyset$, model $\mathcal{M}|^\pi$ is defined as $\langle S^{\mathcal{M}}, R^{\mathcal{M}}, Pr^{\mathcal{M}}, I^{\mathcal{M}}|^\pi, B^{\mathcal{M}}|^\pi, V^{\mathcal{M}} \rangle$.*

We use this definition of an updated model in the semantics of actions and the linear inequalities of probabilities. The rest of the semantics is far more straightforward.

Definition 2.2.9 (Semantics) *Given pointed model $\mathcal{M}, s$, the truth relation is defined as follows:*

$$\begin{array}{rcl}
\mathcal{M}, s \vDash p & \iff & s \in V^{\mathcal{M}}(p) \\
\mathcal{M}, s \vDash \neg\varphi & \iff & \mathcal{M}, s \nvDash \varphi \\
\mathcal{M}, s \vDash \varphi \wedge \psi & \iff & \mathcal{M}, s \vDash \varphi \text{ and } \mathcal{M}, s \vDash \psi \\
\mathcal{M}, s \vDash [a]\varphi & \iff & \text{for all } s' : s \xrightarrow{a} s' \text{ implies } \mathcal{M}|^a, s' \vDash \varphi \\
\mathcal{M}, s \vDash \sum_{i=1}^n q_i B_{\pi_i}\varphi_i \geq q & \iff & \sum_{i=1}^n q_i \mu_{\pi_i}^{\mathcal{M}}([\![\varphi_i]\!]^{\mathcal{M}|^{\pi_i}}) \geq q
\end{array}$$

where $[\![\varphi]\!]^{\mathcal{M}|^{\pi_i}} = \{s \in I^{\mathcal{M}|^{\pi_i}} \mid \mathcal{M}|^{\pi_i}, s \vDash \varphi\}$.

Remark 2.2.10 *Note that if $\mathcal{M}, s$ is a pointed model, i.e., $s \in I^{\mathcal{M}}$, and $s \xrightarrow{a} s'$ then $\mathcal{M}|^a, s'$ is also a pointed model, i.e., $s' \in I^{\mathcal{M}}|^a = I^{\mathcal{M}|^a}$.*

Proposition 2.2.11 *The function $\mu_\pi^{\mathcal{M}}$ is a non-normalized probability, and it has the following properties:*

(1) $\mu_\pi^{\mathcal{M}}([\![\varphi]\!]^{\mathcal{M}|^\pi}) \geq 0$;

(2) $\mu_\pi^{\mathcal{M}}([\![\varphi]\!]^{\mathcal{M}|^\pi}) + \mu_\pi^{\mathcal{M}}([\![\neg\varphi]\!]^{\mathcal{M}|^\pi}) = \mu_\pi^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^\pi})$

(3) $\mu_\epsilon^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}}) = 1$

(4) $\mu_{\pi a}^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^{\pi a}}) = \mu_\pi^{\mathcal{M}}([\![\langle a\rangle\top]\!]^{\mathcal{M}|^\pi})$

PROOF Since $[\![\varphi]\!]^{\mathcal{M}|^\pi} \subseteq I^{\mathcal{M}}|^\pi$, (1) is obvious by Definition 2.2.4. Since $[\![\neg\varphi]\!]^{\mathcal{M}|^\pi} = I^{\mathcal{M}}|^\pi \setminus [\![\varphi]\!]^{\mathcal{M}|^\pi}$ and $[\![\top]\!]^{\mathcal{M}|^\pi} = I^{\mathcal{M}}|^\pi$, (2) is obvious by Definition 2.2.4. Since $\mu_\epsilon^{\mathcal{M}} = B^{\mathcal{M}}$, (3) is obvious. For (4), let $\pi = a_1 \cdots a_n$ and $a_{n+1} = a$ then we have the following:

$$
\begin{aligned}
&\mu_{\pi a}^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^{\pi a}})\\
=&\mu_{\pi a}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi a})\\
=&\sum_{s_0\cdots s_{n+1}\in EP_{\mathcal{M}}(\pi a)} (B^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n+1} Pr^{\mathcal{M}}(s_{i-1}, a_i, s_i))\\
=&\sum_{\{s_0\cdots s_n\in EP_{\mathcal{M}}(\pi)\mid\exists t:t\in R_a^{\mathcal{M}}(s_n)\}}\\
&(B^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}}(s_{i-1}, a_i, s_i) \times (\sum_{t\in R_a^{\mathcal{M}}(s_n)} Pr^{\mathcal{M}}(s_n, a, t)))\\
=&\sum_{\{s_0\cdots s_n\in EP_{\mathcal{M}}(\pi)\mid\exists t:t\in R_a^{\mathcal{M}}(s_n)\}} (B^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}}(s_{i-1}, a_i, s_i))\\
=&\sum_{\{s_0\cdots s_n\in EP_{\mathcal{M}}(\pi)\mid s_n\in[\![\langle a\rangle\top]\!]^{\mathcal{M}|^\pi}\}} (B^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}}(s_{i-1}, a_i, s_i))\\
=&\mu_\pi^{\mathcal{M}}([\![\langle a\rangle\top]\!]^{\mathcal{M}|^\pi})
\end{aligned}
$$

□

Proposition 2.2.12 *Given $\pi = a_1 \cdots a_n$, we have that $\mu_\pi^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^\pi}) = 1$ if and only if $\mathcal{M}, s \vDash K(\!|\pi|\!)\top$.*[4]

PROOF Let $\pi_{(i)} = a_1 \cdots a_i$ for each $1 \leq i \leq n$ and $\pi_{(0)} = \epsilon$ then it is easy to show that $\mathcal{M}, s \vDash K(\!|\pi|\!)\top$ if and only if $\mathcal{M}|^{\pi_{(i)}}, v \vDash \langle a_{i+1}\rangle\top$ for each $0 \leq i < n$ and each $v \in I^{\mathcal{M}}|^{\pi_{(i)}}$.

From left to right: It follows by Proposition 2.2.11 that for each $0 \leq i < n$, we have

$$\mu_{\pi_{(i+1)}}^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^{\pi_{(i+1)}}}) = \mu_{\pi_{(i)}}^{\mathcal{M}}([\![\langle a_{i+1}\rangle\top]\!]^{\mathcal{M}|^{\pi_{(i)}}}) \leq \mu_{\pi_{(i)}}^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^{\pi_{(i)}}}).$$

Since $\mu_\pi^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^\pi}) = 1$ and $\mu_\epsilon^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}}) = 1$, it follows that $\mu_{\pi_{(i)}}^{\mathcal{M}}([\![\langle a_{i+1}\rangle\top]\!]^{\mathcal{M}|^{\pi_{(i)}}}) = \mu_{\pi_{(i)}}^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^{\pi_{(i)}}})$ for each $0 \leq i < n$. Assume that $\mathcal{M}, s \nvDash K(\!|\pi|\!)\top$ then it follows that there are $0 \leq j < n$ and $v \in I^{\mathcal{M}}|^{\pi_{(j)}}$ such that $\mathcal{M}, v \nvDash \langle a_{j+1}\rangle\top$. Since $v \in I^{\mathcal{M}}|^{\pi_{(j)}}$,

[4] Please recall that $(\!|a_1 \cdots a_n|\!)\top := (\!|a_1|\!) \cdots (\!|a_n|\!)\top$ and $(\!|a|\!)\varphi := \langle a\rangle\varphi \wedge [a]\varphi$.

it follows by Definition 2.2.4 that $\mu^{\mathcal{M}}_{\pi_{(j)}} > 0$. Thus, we have $\mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\langle a_{j+1}\rangle\top\rrbracket^{\mathcal{M}|^{\pi(j)}}) < \mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\top\rrbracket^{\mathcal{M}|^{\pi(j)}})$. Contradiction. Therefore, we have $\mathcal{M}, s \vDash K(\!|\pi|\!)\top$.

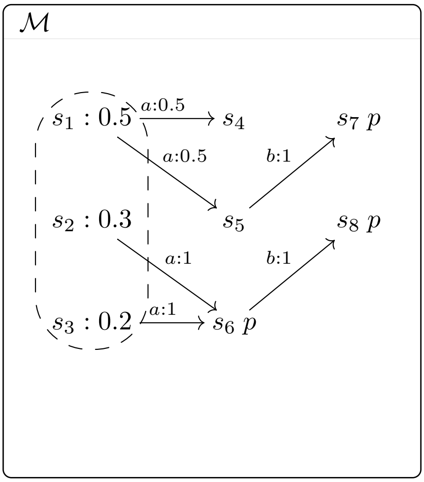
From right to left: We will show that $\mu^{\mathcal{M}}_{\pi_{(i)}}(\llbracket\top\rrbracket^{\mathcal{M}|^{\pi(i)}}) = 1$ for each $0 \leq i \leq n$ by induction on i. It is obvious if $i = 0$. If $i = j+1$ where $0 \leq j < n$, $\mu^{\mathcal{M}}_{\pi_{(j+1)}} = \mu^{\mathcal{M}}_{\pi_{(j)}a_{j+1}}$. It follows by Proposition 2.2.11 that we have the following:

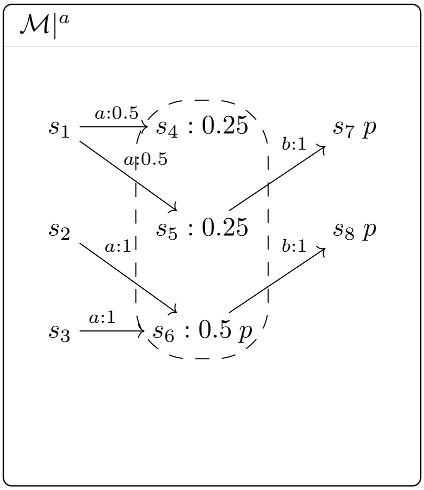
$$\mu^{\mathcal{M}}_{\pi_{(j)}a_{j+1}}(\llbracket\top\rrbracket^{\mathcal{M}|^{\pi(j)a_{j+1}}}) = \mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\langle a_{j+1}\rangle\top\rrbracket^{\mathcal{M}|^{\pi(j)}})$$
$$\mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\langle a_{j+1}\rangle\top\rrbracket^{\mathcal{M}|^{\pi(j)}}) + \mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\neg\langle a_{j+1}\rangle\top\rrbracket^{\mathcal{M}|^{\pi(j)}}) = \mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\top\rrbracket^{\mathcal{M}|^{\pi(j)}})$$

It follows by IH that $\mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\top\rrbracket^{\mathcal{M}|^{\pi(j)}}) = 1$. Therefore, we only need to show that $\mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\neg\langle a_{j+1}\rangle\top\rrbracket^{\mathcal{M}|^{\pi(j)}}) = 0$. Since $\mathcal{M}, s \vDash K(\!|\pi|\!)\top$, namely, $\mathcal{M}|^{\pi(i)}, v \vDash \langle a_{i+1}\rangle\top$ for each $0 \leq i < n$ and each $v \in I^{\mathcal{M}}|^{\pi(i)}$, we have $\llbracket\neg\langle a_{j+1}\rangle\top\rrbracket^{\mathcal{M}|^{\pi(j)}} = \emptyset$. Thus, $\mu^{\mathcal{M}}_{\pi_{(j)}}(\llbracket\neg\langle a_{j+1}\rangle\top\rrbracket^{\mathcal{M}|^{\pi(j)}}) = 0$. □

Recall that $[a]\varphi$ means that φ holds after executing a, and $\mu^{\mathcal{M}}_{\pi_i}(\llbracket\varphi_i\rrbracket^{\mathcal{M}|^{\pi_i}})$ is the probability of reaching φ_i by executing π_i. We will show how the semantics works by working through an example.

Example 2.2.13





1. $\mathcal{M}, s_1 \vDash B_\epsilon(\!|a|\!)\top = 1$
2. $\mathcal{M}, s_1 \vDash [a]B_\epsilon p = 0.5$
3. $\mathcal{M}, s_1 \vDash B_a p = 0.5$
4. $\mathcal{M}, s_1 \vDash B_\epsilon(\!|a|\!)(\!|b|\!)\top = 0.5$
5. $\mathcal{M}, s_1 \vDash [ab]B_\epsilon p = 1$
6. $\mathcal{M}, s_1 \vDash B_{ab} p = 0.75$

1. *This formula shows that initially action a is executable in the set I (where I is the set of worlds the agent considers possible, i.e. $\{s_1, s_2, s_3\}$ indicated by the dotted line). In this sense, the agent knows or is certain that a is executable.*

2. *This formula shows that after all executions of action a the agent assigns probability 0.5 to the set of p-states (i.e. the singleton state s_6).*

3. *This formula says that the agent assigns probability 0.5 to end up in a* p*-state by successfully executing* a*. Remember this is the probability of the executability of* a *that we found above (which was 1) multiplied with the conditional probability of* p *given that* a *was executed, which we also found above (which was 0.5). Of course, the result is also 0.5.*

4. *This formula expresses that initially the sequence of actions* a *and then* b *is executable with probability 0.5, because the formula is only true in* s_2 *and* s_3*.*

5. *This formula expresses that after a successful execution of the sequence* ab *it is certain that* p*, because* p *is true in both* s_7 *and* s_8*.*

6. *This formula expresses that the probability of ending up in a* p*-state by successfully executing* ab *is 0.75. In contrast to the formula in* 4 *we now also take into account that* s_1, s_5, s_7 *is an* ab *execution path.*

We will use this logic as a tool to develop a framework for probabilistic conformant planning, which we can now define in a precise way.

Definition 2.2.14 (Conformant Probabilistic planning) *Given a model* $\mathcal{M}, s$*, a goal formula* φ*, and a threshold* δ*, probabilistic conformant planning for* φ *over* $\mathcal{M}, s$ *w.r.t.* δ *is to find a linear plan* $\pi \in \mathbf{Act}^*$ *such that* $\mathcal{M}, s \vDash B_\pi\varphi \geq \delta$*, where* π *is called a solution to the probabilistic planning problem.*

According to the above definition, to verify that π is a solution is to model check $B_\pi\varphi \geq \delta$ in the pointed model. In the above example, according to item 6, if $\delta \leq 0.75$ then ab is a solution to the probabilistic planning problem for p over $\mathcal{M}, s_1$ w.r.t. δ.

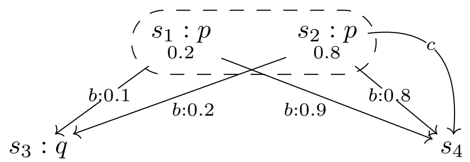
Proposition 2.2.15 *Given* $\mathcal{M}, s$ *and* φ*, if* $\delta = 1$ *then the probabilistic conformant planning problem for a non-probabilistic* φ *over* $\mathcal{M}, s$ *w.r.t.* δ *is a standard conformant planning problem for* φ *over* $\mathcal{M}, s$ *where the probabilities over the states and transitions do not matter, i.e.* $\mathcal{M}, s \vDash B_\pi\varphi = 1 \iff \mathcal{M}, s \vDash K(\!|\pi|\!)\varphi$*.*

PROOF We only need to show that $\mu_\pi^{\mathcal{M}}([\![\varphi]\!]^{\mathcal{M}|^\pi}) = 1$ if and only if $\mathcal{M}, s \vDash K(\!|\pi|\!)\varphi$.

From left to right: Since $[\![\varphi]\!]^{\mathcal{M}|^\pi} \subseteq [\![\top]\!]^{\mathcal{M}|^\pi}$, it follows that $\mu_\pi^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^\pi}) \geq 1$. It follows by Proposition 2.2.11 that $\mu_\pi^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^\pi}) \leq \mu_\epsilon^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}}) = 1$. Therefore, $\mu_\pi^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^\pi}) = 1$. It follows by Proposition 2.2.12 that $\mathcal{M}, s \vDash K(\!|\pi|\!)\top$. Thus, we only need to show that $[\![\neg\varphi]\!]^{\mathcal{M}|^\pi} = \emptyset$. Since $\mu_\pi^{\mathcal{M}}([\![\varphi]\!]^{\mathcal{M}|^\pi}) = \mu_\pi^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^\pi}) = 1$, it follows by Proposition 2.2.11 that $\mu_\pi^{\mathcal{M}}([\![\neg\varphi]\!]^{\mathcal{M}|^\pi}) = 0$. Therefore, $[\![\neg\varphi]\!]^{\mathcal{M}|^\pi} = \emptyset$.

From right to left: Since $\mathcal{M}, s \vDash K(\!|\pi|\!)\varphi$, we have that $\mathcal{M}, s \vDash K(\!|\pi|\!)\top$ and $[\![\neg\varphi]\!]^{\mathcal{M}|^\pi} = \emptyset$. It follows by Proposition 2.2.12 that $\mu_\pi^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}|^\pi}) = 1$. Since $[\![\neg\varphi]\!]^{\mathcal{M}|^\pi} = \emptyset$, we have $\mu_\pi^{\mathcal{M}}([\![\neg\varphi]\!]^{\mathcal{M}|^\pi}) = 0$. It follows by Proposition 2.2.11 that $\mu_\pi^{\mathcal{M}}([\![\varphi]\!]^{\mathcal{M}|^\pi}) = 1$.
□

Example 2.2.16 (The pill or surgery problem) *Please consider the following scenario. There is a disease (let* p *signify that the agent has the disease), which the agent certainly has. There are two kinds of treatment available: surgery or pills. Surgery will very likely be effective, but unfortunately surgery comes with the risk that unpleasant side-effects may occur (let* q *express that these side-effects occur). Pills will certainly work, but* *only for certain genotypes of agents. For agents with the wrong genotype, allergies will prevent the agent being able to swallow the pills. The agent believes with probability* 0.8 *that she has the right genotype. If she has the right genotype, surgery will not be as likely to succeed as when she does not have the right genotype. Let's say that if she has the wrong genotype surgery will succeed without any side-effects with probability* 0.9 *and have side-effects with probability* 0.1*. If she has the right genotype, surgery will succeed without side-effects with probability* 0.8 *and will lead to side-effects with probability* 0.2*. This situation is displayed below. Here, action* b *is having surgery, and action* c *is taking pills. Note that not swallowing the pills means that action* c *is not executable (because the agent will not be able to really take the pills).*



$B_b(\neg p \wedge \neg q) = 0.8 \times 0.8 + 0.2 \times 0.9 = 0.82$
$B_c(\neg p \wedge \neg q) = 0.8$

Suppose that the goal of the agent is to become healthy with no side-effects. Let's say that the threshold is 0.81. Which course of action is best? As is clear from the calculation above, plan b, i.e. surgery, will yield the desired result with probability 0.82, whereas plan c will lead to the desired goal with probability 0.8, and so it is best for the agent to have surgery.

2.3 A deductive system

2.3.1 Axiom system $\mathbb{SLCPP}$

In this section, we provide a Hilbert-style proof system for the logic presented above. A proof consists of a sequence of formulas such that each formula is either an instance of an axiom or it can be obtained by applying one of the rules to formulas occurring earlier in the sequence.

Definition 2.3.1 ($\mathbb{SLCPP}$ System) *The axiom system $\mathbb{SLCPP}$ is shown in Table 2.1. We write $\mathbb{SLCPP} \vdash \varphi$ (or sometimes just $\vdash \varphi$) to mean that the formula φ is* derivable *in the axiomatic system $\mathbb{SLCPP}$; the negation of $\mathbb{SLCPP} \vdash \varphi$ is written $\mathbb{SLCPP} \nvdash \varphi$ (or just $\nvdash \varphi$). To say that a set D of formulas is $\mathbb{SLCPP}$-*inconsistent *(or just inconsistent) means that there is a finite subset $D' \subseteq D$ such that $\vdash \neg \bigwedge D'$, where $\bigwedge D' := \bigwedge_{\varphi \in D'} \varphi$ if $D' \neq \emptyset$ and $\bigwedge_{\varphi \in \emptyset} \varphi := \top$. To say that a set of formulas is $\mathbb{SLCPP}$-*consistent *(or just consistent) means that the set of formulas is not inconsistent. Consistency or inconsistency of a formula refers to the consistency or inconsistency of the singleton set containing the formula.*

The linear inequality axioms can be found in Definition A.2.4 of Appendix A.2.1. Let us explain how the above axioms are to be read. We only focus on those involving probability. Axiom `T` expresses that truths are assigned a positive probability. This is because the empty sequence is always executable and we defined pointed models such that the state is always in $I^{\mathcal{M}}$, so it will always receive positive probability.

	AXIOMS
	All instances of propositional tautologies
	All instances of linear inequality axioms
`DIST`(a)	$[a](\varphi \to \psi) \to ([a]\varphi \to [a]\psi)$
`T`	$K\varphi \to \varphi$
`Nonneg`(π)	$B_\pi\varphi \geq 0$
`PRTR`(ϵ)	$K\top$
`PRF`(πa)	$B_{\pi a}\varphi \leq B_\pi\langle a\rangle\varphi$
`PRFEQ`(πa)	$B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) = 0 \leftrightarrow B_{\pi a}\varphi = B_\pi\langle a\rangle\varphi$
`Add`(π)	$B_\pi(\varphi \wedge \psi) + B_\pi(\varphi \wedge \neg\psi) = B_\pi\varphi$
`ITSP`	$(\sum_{i=1}^n q_i B_{\pi\pi'_i}\varphi_i \bowtie qB_\pi\top) \to B_\pi(\sum_{i=1}^n q_i B_{\pi'_i}\varphi_i \bowtie q) = B_\pi\top$
`CP`	$\langle a\rangle\top \to ([a](\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \bowtie q) \leftrightarrow \sum_{i=1}^n q_i B_{a\pi_i}\varphi_i \bowtie qB_a\top)$
`DET`	$\langle a\rangle\varphi \to [a]\varphi$ where φ is a probability formula
	RULES
MP	From $\varphi \to \psi$ and φ, infer ψ
GEN	From φ, infer $[a]\varphi$
Equivalence	From $\varphi \leftrightarrow \psi$, infer $B_\pi\varphi = B_\pi\psi$

Table 2.1: System $\mathbb{SLCPP}$
(Please recall that $\bowtie$ is one of $\leq, \geq, <, >, =, \neq$.)

Axiom `Nonneg`(π) expresses that any formula receives a non-negative probability (since negative probabilities don't make sense).

Axiom `PRTR`(ϵ) expresses that the set of states that the agent considers possible is assigned probability 1.

Axiom `PRF`(πa) expresses that the probability of those πa-execution paths leading to φ states, is less than or equal to the probability of those π-execution paths leading to states where a can lead to a φ state. This is because executing π may lead to a state where executing a may lead to a φ-state, but executing a could also lead to a non-φ-state.

Axiom `PRFEQ`(πa) expresses the condition under which the above probabilities are equal. This is the case if either all a-paths in π-reachable states lead to φ-states or if all a-paths in π-reachable states lead to non-φ-states, or in other words whenever the probability that executing a can lead to a φ-state and can lead to a non-φ-state is zero.

Axiom `Add`(π) expresses that probabilities are additive.

Axiom `ITSP` is the combination of `4` and `5`. Two simple forms of `ITSP` are the following.

$$B_\epsilon\varphi \geq q \to B_\epsilon(B_\epsilon\varphi \geq q) = 1$$
$$\neg(B_\epsilon\varphi \geq q) \to B_\epsilon(B_\epsilon\varphi < q) = 1$$

Axiom `CP` is essentially the definition of the update using normalization, given that a is executable. A simple form of `CP` is the following.

$$\langle a\rangle\top \to ([a](B_\epsilon\varphi \bowtie q) \leftrightarrow B_a\varphi \bowtie qB_a\top)$$

Note that DET is not valid for arbitrary ψ. It is crucial that a is not deterministic for basic facts.

Proposition 2.3.2 $\vdash B_\pi\bot = 0$

PROOF It follows by Axiom $\mathtt{Add}(\pi)$ that $\vdash B_\pi\top + B_\pi\bot = B_\pi\top$, namely $\vdash B_\pi\top + B_\pi\bot + (-1)B_\pi\top = 0$. It follows by the axioms of linear inequality logic, addition of coefficients, that $\vdash B_\pi\bot + (0)B_\pi\top = 0$. By the 0-term axiom of linear inequality logic, it follows that $\vdash B_\pi\bot = 0$. □

Proposition 2.3.3 $\vdash B_\pi\varphi = B_\pi\top \to B_\pi\neg\varphi = 0$

PROOF It follows by Axiom $\mathtt{Add}(\pi)$ that $\vdash B_\pi\varphi + B_\pi\neg\varphi = B_\pi\top$. Therefore, by linear inequality logic, we have $\vdash B_\pi\varphi = B_\pi\top \to B_\pi\neg\varphi = 0$. □

Proposition 2.3.4 *If* $\vdash \psi \leftrightarrow \chi$ *then* $\vdash \varphi \leftrightarrow \varphi(\psi/\chi)$.

PROOF We prove it by induction on φ. We only focus on the case of $\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \geq q$; the other cases are straightforward.

If $\varphi := \sum_{i=1}^n q_i B_{\pi_i}\varphi_i \geq q$, it follows by IH that $\vdash \varphi_i \leftrightarrow \varphi_i(\psi/\chi)$ for each $1 \leq i \leq n$. By Rule Equivalence presented in Table 2.1, we have that $B_{\pi_i}\varphi_i = B_{\pi_i}\varphi_i(\psi/\chi)$. It follows by linear inequality logic that $\vdash \varphi \leftrightarrow \sum_{i=1}^n q_i B_{\pi_i}\varphi_i(\psi/\chi) \geq q$. □

Proposition 2.3.5 *If* $\vdash \varphi \to \psi$ *then we have* $\vdash B_\pi\varphi \leq B_\pi\psi$.

PROOF It follows by $\vdash \varphi \to \psi$ that $\vdash \varphi \vee \psi \leftrightarrow \psi$. By the Equivalence rule, we have $\vdash B_\pi(\varphi \vee \psi) = B_\pi\psi$. It follows by Axiom $\mathtt{Add}(\pi)$ that $\vdash B_\pi((\varphi \vee \psi) \wedge \varphi) + B_\pi((\varphi \vee \psi) \wedge \neg\varphi) = B_\pi(\varphi \vee \psi)$, and then $\vdash B_\pi((\varphi \vee \psi) \wedge \varphi) + B_\pi((\varphi \vee \psi) \wedge \neg\varphi) = B_\pi\psi$. Since $\vdash (\varphi \vee \psi) \wedge \varphi \leftrightarrow \varphi$, it follows $\vdash B_\pi((\varphi \vee \psi) \wedge \varphi) = B_\pi\varphi$. Thus, we have $\vdash B_\pi\varphi + B_\pi((\varphi \vee \psi) \wedge \neg\varphi) = B_\pi\psi$. It follows by Axiom $\mathtt{Nonneg}(\pi)$ that $\vdash B_\pi((\varphi \vee \psi) \wedge \neg\varphi) \geq 0$. Thus, we have $\vdash B_\pi\varphi \leq B_\pi\psi$. □

Proposition 2.3.6 $\vdash B_{\pi a}\top = B_\pi\langle a\rangle\top$

PROOF It follows by MP and GEN that $\vdash \langle a\rangle\bot \leftrightarrow \bot$. Thus we have $\vdash (\langle a\rangle\top \wedge \langle a\rangle\bot) \leftrightarrow \bot$. It follows by rule Equivalence that $B_\pi(\langle a\rangle\top \wedge \langle a\rangle\bot) = B_\pi\bot$. It follows by Proposition 2.3.2 that $B_\pi(\langle a\rangle\top \wedge \langle a\rangle\bot) = 0$. It follows by Axiom $\mathtt{PRFEQ}(\pi a)$ that $\vdash B_{\pi a}\top = B_\pi\langle a\rangle\top$. □

Proposition 2.3.7 $\vdash B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) > 0 \leftrightarrow B_{\pi a}\varphi < B_\pi\langle a\rangle\varphi$

PROOF (1) $\vdash B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) \neq 0 \leftrightarrow B_{\pi a}\varphi \neq B_\pi\langle a\rangle\varphi$ by Axiom $\mathtt{PRFEQ}(\pi a)$
(2) $\vdash B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) \neq 0 \leftrightarrow B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) > 0$ by Axiom $\mathtt{Nonneg}(\pi)$
(3) $B_{\pi a}\varphi \neq B_\pi\langle a\rangle\varphi \leftrightarrow B_{\pi a}\varphi < B_\pi\langle a\rangle\varphi$ by Axiom $\mathtt{PRF}(\pi a)$
(4) $\vdash B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) > 0 \leftrightarrow B_{\pi a}\varphi < B_\pi\langle a\rangle\varphi$ by (1)-(3) □

Proposition 2.3.8 $\vdash [a](\sum_{i=1}^{n} q_i B_{\pi_i}\varphi_i \bowtie q \vee \psi) \leftrightarrow (\sum_{i=1}^{n} q_i B_{a\pi_i}\varphi_i \bowtie qB_a\top) \vee [a]\psi$

PROOF Firstly, to make the proof shorter, let $\chi := \sum_{i=1}^{n} q_i B_{\pi_i}\varphi_i \bowtie q$ and $\chi' := \sum_{i=1}^{n} q_i B_{a\pi_i}\varphi_i \bowtie q$.
$\Rightarrow$
(1) $\vdash [a](\chi \vee \psi) \wedge [a]\neg\chi \to [a]\psi$ by normal modal logic
(2) $\vdash \langle a\rangle\chi \to \langle a\rangle\top \wedge [a]\chi$ by normal modal logic and Axiom DET
(3) $\vdash \langle a\rangle\top \wedge [a]\chi \to \chi'$ by Axiom CP
(4) $\vdash \langle a\rangle\chi \to \chi'$ by (2), (3)
(5) $\vdash \neg\chi' \to [a]\neg\chi$ by (4)
(6) $\vdash [a](\chi \vee \psi) \wedge \neg\chi' \to [a]\psi$ by (1) and (5)
(7) $\vdash [a](\chi \vee \psi) \to \chi' \vee [a]\psi$ by (7)
$\Leftarrow$
(1) $\vdash [a]\psi \to [a](\chi \vee \psi)$ by normal modal logic
(2) $\vdash \chi' \to (\langle a\rangle\top \to [a]\chi)$ by Axiom CP
(3) $\vdash \chi' \to [a]\bot \vee [a]\chi$ by (2)
(4) $\vdash [a]\bot \to [a]\chi$ by normal modal logic
(5) $\vdash \chi' \to [a]\chi$ by (3) and (4)
(6) $\vdash \chi' \to [a](\chi \vee \psi)$ by (5)
(7) $\vdash \chi' \vee [a]\psi \to [a](\chi \vee \psi)$ by (1) and (6) $\square$

Proposition 2.3.9 $\vdash \neg(\sum_{i=1}^{n} q_i B_{\pi\pi_i'}\varphi_i \bowtie qB_\pi\top) \to B_\pi(\sum_{i=1}^{n} q_i B_{\pi_i'}\varphi_i \bowtie q) = 0$

PROOF Let χ be the formula $\sum_{i=1}^{n} q_i B_{\pi_i'}\varphi_i \bowtie q$, and let $\overline{\bowtie}$ denote $>, <, \geq, \leq, \neq$ or $=$ if $\bowtie$ is $\leq, \geq, <, >, =$, or $\neq$, respectively. Please note that $\neg\chi := \sum_{i=1}^{n} q_i B_{\pi_i'}\varphi_i \overline{\bowtie} q$ and $\neg(\sum_{i=1}^{n} q_i B_{\pi\pi_i'}\varphi_i \bowtie qB_\pi\top) := \sum_{i=1}^{n} q_i B_{\pi\pi_i'}\varphi_i \overline{\bowtie} qK_\pi\top$. Then, we have the following:
(1) $\vdash \neg(\sum_{i=1}^{n} q_i B_{\pi\pi_i'}\varphi_i \bowtie qB_\pi\top) \to B_\pi\neg\chi = B_\pi\top$, by Axiom ITSP.
(2) $\vdash B_\pi\chi + B_\pi\neg\chi = B_\pi\top$, by Axiom Add($\pi$)
(3) $\vdash B_\pi\neg\chi = B_\pi\top \to B_\pi\chi = 0$, by (2) and Linear Inequality Logic
(4) $\vdash \neg(\sum_{i=1}^{n} q_i B_{\pi\pi_i'}\varphi_i \bowtie qB_\pi\top) \to B_\pi\chi = 0$, by (1) and (3) $\square$

Proposition 2.3.10 $\vdash B_\pi\varphi = B_\pi\top \to B_\pi((\varphi \vee \psi) \wedge \chi) = B_\pi\chi$

PROOF (1) $\vdash (B_\pi\varphi \leq B_\pi(\varphi \vee \psi)) \wedge (B_\pi(\varphi \vee \psi) \leq B_\pi\top)$ by Proposition 2.3.6
(2) $\vdash B_\pi\varphi = B_\pi\top \to B_\pi\top \leq B_\pi(\varphi \vee \psi)$ by (1) and linear inequality logic
(3) $\vdash B_\pi\varphi = B_\pi\top \to B_\pi(\varphi \vee \psi) = B_\pi\top$ by (1) and (2)
(4) $\vdash B_\pi(\varphi \vee \psi) + B_\pi\neg(\varphi \vee \psi) = B_\pi\top$ by Axiom Add(π)
(5) $\vdash B_\pi(\varphi \vee \psi) = B_\pi\top \to B_\pi\neg(\varphi \vee \psi) = 0$ by (4)
(6) $\vdash B_\pi(\neg(\varphi \vee \psi) \wedge \chi) \leq B_\pi\neg(\varphi \vee \psi)$ by Proposition 2.3.6
(7) $\vdash B_\pi(\varphi \vee \psi) = B_\pi\top \to B_\pi(\neg(\varphi \vee \psi) \wedge \chi) \leq 0$ by (5) and (6)
(8) $\vdash B_\pi(\varphi \vee \psi) = B_\pi\top \to B_\pi(\neg(\varphi \vee \psi) \wedge \chi) = 0$ by (7) and Axiom Nonneg(π)
(9) $\vdash B_\pi((\varphi \vee \psi) \wedge \chi) + B_\pi(\neg(\varphi \vee \psi) \wedge \chi) = B_\pi\chi$ by Axiom Add(π)
(10) $\vdash B_\pi(\varphi \vee \psi) = B_\pi\top \to B_\pi((\varphi \vee \psi) \wedge \chi) = B_\pi\chi$ by (8) and (9)
(11) $\vdash B_\pi\varphi = B_\pi\top \to B_\pi((\varphi \vee \psi) \wedge \chi) = B_\pi\chi$ by (3) and (10) $\square$

Proposition 2.3.11 $\vdash B_\pi\varphi = 0 \to B_\pi(\varphi \wedge \chi) = 0$

PROOF (1) $\vdash B_\pi(\varphi \wedge \chi) \leq B_\pi\varphi$ by Proposition 2.3.6
(2) $\vdash B_\pi\varphi = 0 \to B_\pi(\varphi \wedge \chi) \leq 0$ by (1)
(3) $\vdash B_\pi\varphi = 0 \to B_\pi(\varphi \wedge \chi) = 0$ by Axiom `Nonneg`(π) and (2) □

Proposition 2.3.12 $\vdash B_\pi\varphi = 0 \to B_\pi((\varphi \vee \psi) \wedge \chi) = B_\pi(\psi \wedge \chi)$

PROOF Let $\delta_0 := \varphi \wedge \chi$ and $\delta_1 := \psi \wedge \chi$. We have the followings:
(1) $B_\pi((\varphi \vee \psi) \wedge \chi) = B_\pi(\delta_0 \vee \delta_1)$ by Rule Equivalence
(2) $\vdash B_\pi\varphi = 0 \to B_\pi\delta_0 = 0$ by Proposition 2.3.11
(3) $\vdash B_\pi\delta_1 = B_\pi((\delta_0 \vee \delta_1) \wedge \delta_1)$ by Rule Equivalence
(4) $\vdash B_\pi(\delta_0 \wedge \neg\delta_1) = B_\pi((\delta_0 \vee \delta_1) \wedge \neg\delta_1)$ by Rule Equivalence
(5) $\vdash B_\pi\delta_0 = 0 \to B_\pi(\delta_0 \wedge \neg\delta_1) = 0$ by Proposition 2.3.11
(6) $\vdash B_\pi\varphi = 0 \to B_\pi(\delta_0 \wedge \neg\delta_1) = 0$ by (2) and (5)
(7) $\vdash B_\pi\varphi = 0 \to B_\pi((\delta_0 \vee \delta_1) \wedge \neg\delta_1) = 0$ by (4) and (6)
(8) $\vdash B_\varphi((\delta_0 \vee \delta_1) \wedge \delta_1) + B_\pi((\delta_0 \vee \delta_1) \wedge \neg\delta_1 = B_\pi(\delta_0 \vee \delta_1)$ by Axiom `Add`(π)
(9) $\vdash B_\pi\varphi = 0 \to B_\pi(\delta_0 \vee \delta_1) = B_\pi\delta_1$ by (3), (7) and (8)
(10) $\vdash B_\pi\varphi = 0 \to B_\pi((\varphi \vee \psi) \wedge \chi) = B_\pi\delta_1$ by (1) and (9) □

Proposition 2.3.13 *Let* $T := \sum_{j=1}^m q_j B_{\pi_j}\psi_j$, $\delta_0 := \sum_{i=1}^n q'_i B_{\pi'_i}\varphi_i \bowtie_1 q$, *and* $\delta_1 := \sum_{i=1}^n q'_i B_{\pi'_i}\varphi_i \bowtie_1 q B_\pi\top$. *We have* $\vdash q_0 B_\pi((\delta_0 \vee \psi) \wedge \chi) + T \geq q \leftrightarrow (\delta_1 \wedge (q_0 B_\pi\chi + T \bowtie_2 q)) \vee (\neg\delta_1 \wedge (q_0 B_\pi(\psi \wedge \chi) + T \bowtie_2 q))$.

PROOF (1) $\vdash \delta_1 \to B_\pi\delta_0 = B_\pi\top$ by Axiom `ITSP`
(2) $\vdash B_\pi\delta_0 = B_\pi\top \to B_\pi((\delta_0 \vee \psi) \wedge \chi) = B_\pi\chi$ by Proposition 2.3.10
(3) $\vdash \delta_1 \to B_\pi((\delta_0 \vee \psi) \wedge \chi) = B_\pi\chi$ by (1) and (2)
(4) $\vdash \neg\delta_1 \to B_\pi\neg\delta_0 = B_\pi\top$ by Axiom `ITSP`
(5) $\vdash B_\pi\neg\delta_0 = B_\pi\top \to B_\pi\delta_0 = 0$ by Proposition 2.3.3
(6) $\vdash \neg\delta_1 \to B_\pi\delta_0 = 0$ by (4) and (5)
(7) $\vdash B_\pi\delta_0 = 0 \to B_\pi((\delta_0 \vee \psi) \wedge \chi) = B_\pi(\psi \wedge \chi)$ by Proposition 2.3.12
(8) $\vdash \neg\delta_1 \to B_\pi((\delta_0 \vee \psi) \wedge \chi) = B_\pi(\psi \wedge \chi)$ by (6) and (7)
(9) $\vdash q_0 B_\pi((\delta_0 \vee \psi) \wedge \chi) + T \bowtie_2 q \leftrightarrow (\delta_1 \wedge (q_0 B_\pi\chi + T \bowtie_2 q)) \vee (\neg\delta_1 \wedge (q_0 B_\pi(\psi \wedge \chi) + T \bowtie_2 q))$ by (3), (8) and linear inequality logic □

Please note that the simplest version of Proposition 2.3.13 is

$$\vdash B_\epsilon((\varphi \vee \psi) \wedge \chi) \bowtie_2 q \leftrightarrow (\varphi \wedge B_\epsilon\chi \bowtie_2 q) \vee (\neg\varphi \wedge B_\epsilon(\psi \wedge \chi) \bowtie_2 q)$$

where $\varphi := \sum_{i=1}^n q_i\varphi_i \bowtie_1 q$. Therefore, we have the following proposition.

Proposition 2.3.14 *If* $\psi := \sum_{i=1}^n q_i\varphi_i \bowtie_1 q$ *then we have*

(i) $\vdash B_\epsilon(\psi \wedge \chi) \bowtie_2 q \leftrightarrow (\psi \wedge B_\epsilon\chi \bowtie_2 q) \vee (\neg\psi \wedge B_\epsilon\bot \bowtie_2 q)$

(ii) $\vdash B_\epsilon(\psi \vee \chi) \bowtie_2 q \leftrightarrow (\psi \wedge B_\epsilon\top \bowtie_2 q) \vee (\neg\psi \wedge B_\epsilon\chi \bowtie_2 q)$

Proposition 2.3.15 $\vdash B_\pi\langle a\rangle\varphi > 0 \to B_{\pi a}\varphi > 0$

PROOF (1) $\vdash B_\pi\langle a\rangle\neg\varphi \leq B_\pi\langle a\rangle\top$ by Proposition 2.3.5 and $\vdash \langle a\rangle\neg\varphi \to \langle a\rangle\top$
(2) $\vdash B_\pi\langle a\rangle\neg\varphi \leq B_{\pi a}\top$ by Proposition 2.3.6 and (1)
(3) $\vdash B_\pi(\langle a\rangle\neg\varphi \wedge \langle a\rangle\varphi) > 0 \to B_{\pi a}\neg\varphi < B_\pi\langle a\rangle\neg\varphi$ by Proposition 2.3.7
(4) $\vdash B_\pi(\langle a\rangle\neg\varphi \wedge \langle a\rangle\varphi) > 0 \to B_{\pi a}\neg\varphi < B_{\pi a}\top$ by (2) and (3)
(5) $\vdash B_{\pi a}\neg\varphi < B_{\pi a}\top \to B_{\pi a}\varphi > 0$ by $\mathtt{Add}(\pi a)$: $\vdash B_{\pi a}\varphi + B_{\pi a}\neg\varphi = B_{\pi a}\top$
(6) $\vdash B_\pi(\langle a\rangle\neg\varphi \wedge \langle a\rangle\varphi) > 0 \to B_{\pi a}\varphi > 0$ by (4) and (5)
(7) $\vdash B_\pi(\langle a\rangle\neg\varphi \wedge \langle a\rangle\varphi) > 0 \to (B_\pi\langle a\rangle\varphi > 0 \to B_{\pi a}\varphi > 0)$ by (6) and propositional logic
(8) $\vdash B_\pi\langle a\rangle\varphi = B_{\pi a}\varphi \to (B_\pi\langle a\rangle\varphi > 0 \to B_{\pi a}\varphi > 0)$ by linear inequality logic
(9) $\vdash B_\pi(\langle a\rangle\neg\varphi \wedge \langle a\rangle\varphi) = 0 \to (B_\pi\langle a\rangle\varphi > 0 \to B_{\pi a}\varphi > 0)$ by (8) and Axiom $\mathtt{PRFEQ}(\pi a)$
(10) $\vdash B_\pi(\langle a\rangle\neg\varphi \wedge \langle a\rangle\varphi) \geq 0 \to (B_\pi\langle a\rangle\varphi > 0 \to B_{\pi a}\varphi > 0)$ by (7) and (9)
(11) $\vdash B_\pi\langle a\rangle\varphi > 0 \to B_{\pi a}\varphi > 0$ by (10) and Axiom $\mathtt{Nonneg}(\pi a)$ □

Proposition 2.3.16 *If $\psi \wedge \chi$ is inconsistent then* $\vdash B_\pi(\psi \vee \chi) = B_\pi\psi + B_\pi\chi$.

PROOF Since $\psi \wedge \chi$ is inconsistent, it follows that $\vdash \psi \wedge \chi \leftrightarrow \bot$, and then $\vdash (\psi\vee\chi)\wedge\varphi \leftrightarrow \psi$. It follows by Axiom $\mathtt{Add}(\pi)$ that $\vdash B_\pi(\psi\vee\chi)\wedge\psi + B_\pi(\psi\vee\chi)\wedge\neg\psi = B_\pi(\psi \vee \chi)$. Since $\vdash (\psi \vee \chi) \wedge \neg\psi \leftrightarrow \chi$, we have $\vdash B_\pi\psi + B_\pi\chi = B_\pi(\psi \vee \chi)$ □

2.3.2 Soundness

In this section we show that the axiomatization presented in the previous section is sound with respect to the semantics provided in Section 2.2. Given that the logic is built on well-understood modal logic, we will not show that the usual modal axioms and rules are sound. Also the part of the axiomatization concerned with linear inequalities is well-understood and we do not show the soundness of that part either. Instead, we will focus on the axioms and rules that deal with the interplay between actions and probability.

The flowing shows that Axiom $\mathtt{Nonneg}(\pi)$ is valid.

Proposition 2.3.17 $\models B_\pi\varphi \geq 0$

PROOF It follows by Definition 2.2.4 that $\mu_\pi^{\mathcal{M}}([\![\varphi]\!]^{\mathcal{M}|^\pi}) \geq 0$ for each model $\mathcal{M}$. □

The flowing shows that Axiom $\mathtt{PRTR}(\epsilon)$ is valid.

Proposition 2.3.18 $\models K\top$

PROOF We only need to show that $\models B_\epsilon\top = 1$. Given a model $\mathcal{M}$, since $[\![\top]\!]^{\mathcal{M}} = I^{\mathcal{M}}$, it follows that $\mu_\epsilon^{\mathcal{M}}([\![\top]\!]^{\mathcal{M}}) = 1$. □

In order to prove the soundness of $\mathtt{PRF}(\pi a)$, we first prove two auxiliary propositions. The first is about the relation between probabilities in a model after an action and probabilities preceding the action.

Proposition 2.3.19 *Given model $\mathcal{M}$ and $I^{\mathcal{M}}|^\pi \neq \emptyset$, we have $\mu_{\pi'}^{\mathcal{M}|^\pi}(t) = \mu_{\pi\pi'}^{\mathcal{M}}(t)/\mu_\pi^{\mathcal{M}}(I^{\mathcal{M}}|^\pi)$ for each $t \in S^{\mathcal{M}}$.*

PROOF

$$
\begin{array}{rl}
& \mu_{\pi'}^{\mathcal{M}|^{\pi}}(t) \\
= & \displaystyle\sum_{\{s_0\cdots s_n \in EP_{\mathcal{M}|^{\pi}}(\pi')|s_n=t\}} (B^{\mathcal{M}|^{\pi}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}|^{\pi}}(s_{i-1}, a_i, s_i)) \\
= & 1/\mu_{\pi}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi}) \displaystyle\sum_{\{s_0\cdots s_n \in EP_{\mathcal{M}|^{\pi}}(\pi')|s_n=t\}} (\mu_{\pi}^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}|^{\pi}}(s_{i-1}, a_i, s_i)) \\
= & 1/\mu_{\pi}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi}) \displaystyle\sum_{\{s_0\cdots s_n \in EP_{\mathcal{M}|^{\pi}}(\pi')|s_n=t\}} \Bigg(\sum_{\{s'_0\cdots s'_m \in EP_{\mathcal{M}}(\pi)|s'_m=s_0\}} \\
& B^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}}(s'_{i-1}, a'_i, s'_i) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}|^{\pi}}(s_{i-1}, a_i, s_i) \Bigg) \\
= & 1/\mu_{\pi}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi}) \displaystyle\sum_{\{u_0\cdots u_{m+n} \in EP_{\mathcal{M}}(\pi\pi')|u_{m+n}=t\}} (\mu_{\pi}^{\mathcal{M}}(u_0) \times \Pi_{i=1}^{m+n} Pr^{\mathcal{M}}(s_{i-1}, a_i, s_i)) \\
= & \mu_{\pi\pi'}^{\mathcal{M}}(t)/\mu_{\pi}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi})
\end{array}
$$

□

Using Proposition 2.3.19, we can prove the second auxiliary proposition that expresses that updating a model with a composed action is the same as updating the model sequentially, first with the one component of the action, then the other component.

Proposition 2.3.20 *Given model $\mathcal{M}$ and $I^{\mathcal{M}}|^{\pi\pi'} \neq \emptyset$, we have $\mathcal{M}|^{\pi}|^{\pi'} = \mathcal{M}|^{\pi\pi'}$.*

PROOF We only need to show that $B^{\mathcal{M}|^{\pi}|^{\pi'}}(t) = B^{\mathcal{M}|^{\pi\pi'}}(t)$ for each $t \in I^{\mathcal{M}}|^{\pi\pi'}$.

$$
\begin{array}{rcll}
B^{\mathcal{M}|^{\pi}|^{\pi'}}(t) & = & \dfrac{\mu_{\pi'}^{\mathcal{M}|^{\pi}}(t)}{\mu_{\pi'}^{\mathcal{M}|^{\pi}}(I^{\mathcal{M}}|^{\pi\pi'})} & \\
& = & \dfrac{\mu_{\pi\pi'}^{\mathcal{M}}(t)/\mu_{\pi}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi}}{\sum_{s\in I^{\mathcal{M}}|^{\pi\pi'}} \mu_{\pi\pi'}^{\mathcal{M}}(s)/\mu_{\pi}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi})} & \text{by Proposition 2.3.19} \\
& = & \dfrac{\mu_{\pi\pi'}^{\mathcal{M}}(t)}{\sum_{s\in I^{\mathcal{M}}|^{\pi\pi'}} \mu_{\pi\pi'}^{\mathcal{M}}(s)} & \\
& = & B^{\mathcal{M}|^{\pi\pi'}}(t) &
\end{array}
$$

□

Using Proposition 2.3.20, we can show the soundness of $\mathtt{PRF}(\pi a)$.

Proposition 2.3.21 $\models B_{\pi a}\varphi \leq B_{\pi}\langle a\rangle\varphi$.

PROOF Given model $\mathcal{M}$, we only need to show that $\mu_{\pi a}^{\mathcal{M}}([\![\varphi]\!]^{\mathcal{M}|^{\pi a}}) \leq \mu_{\pi}^{\mathcal{M}}([\![\langle a\rangle\varphi]\!]^{\mathcal{M}|^{\pi}})$. If $I^{\mathcal{M}}|^{\pi a} = \emptyset$ then $\mu_{\pi a}^{\mathcal{M}}([\![\varphi]\!]^{\mathcal{M}|^{\pi a}}) = 0$. Since $\mu_{\pi}^{\mathcal{M}}([\![\langle a\rangle\varphi]\!]^{\mathcal{M}|^{\pi}}) \geq 0$, it is obvious. If $I^{\mathcal{M}}|^{\pi a} \neq \emptyset$, we have that for each $t \in [\![\varphi]\!]^{\mathcal{M}|^{\pi a}} \subseteq I^{\mathcal{M}}|^{\pi a}$, there exists $s \in I^{\mathcal{M}}|^{\pi}$ such that $s \xrightarrow{a} t$. Moreover, it follows by Definition 2.2.4 that for each $t \in I^{\mathcal{M}}|^{\pi a}$,

$$\mu_{\pi a}^{\mathcal{M}}(t) = \sum_{\{s\in I^{\mathcal{M}}|^{\pi}|s\xrightarrow{a}t\}} \mu_{\pi}^{\mathcal{M}}(s) \times Pr^{\mathcal{M}}(s, a, t)$$

We then have the following:

$$
\begin{array}{ll}
& \mu^{\mathcal{M}}_{\pi a}([\![\varphi]\!]^{\mathcal{M}|^{\pi a}}) \\
= & \sum_{t\in[\![\varphi]\!]^{\mathcal{M}|^{\pi a}}} \mu^{\mathcal{M}}_{\pi a}(t) \\
= & \sum_{t\in[\![\varphi]\!]^{\mathcal{M}|^{\pi a}}} \big(\sum_{\{s\in I^{\mathcal{M}}|^{\pi}|s\xrightarrow{a}t\}} \mu^{\mathcal{M}}_{\pi}(s) \times Pr^{\mathcal{M}}(s,a,t)\big) \\
= & \sum_{\{s\in I^{\mathcal{M}}|^{\pi} \mid \exists t\in[\![\varphi]\!]^{\mathcal{M}|^{\pi a}}:s\xrightarrow{a}t\}} \mu^{\mathcal{M}}_{\pi}(s) \times \big(\sum_{t\in([\![\varphi]\!]^{\mathcal{M}|^{\pi a}}\cap R^{\mathcal{M}}_a(s))} Pr^{\mathcal{M}}(s,a,t)\big) \\
\leq & \sum_{\{s\in I^{\mathcal{M}}|^{\pi} \mid \exists t\in[\![\varphi]\!]^{\mathcal{M}|^{\pi a}}:s\xrightarrow{a}t\}} \mu^{\mathcal{M}}_{\pi}(s) \\
= & \sum_{\{s\in I^{\mathcal{M}}|^{\pi} \mid \exists t\in[\![\varphi]\!]^{\mathcal{M}|^{\pi}|^{a}}:s\xrightarrow{a}t\}} \mu^{\mathcal{M}}_{\pi}(s) \qquad \text{by Proposition 2.3.20} \\
= & \sum_{s\in[\![\langle a\rangle\varphi]\!]^{\mathcal{M}|^{\pi}}} \mu^{\mathcal{M}}_{\pi}(s) \\
= & \mu^{\mathcal{M}}_{\pi}([\![\langle a\rangle\varphi]\!]^{\mathcal{M}|^{\pi}})
\end{array}
$$

□

The soundness of this axiom is used in the proof of the soundness of PRFEQ(πa).

Proposition 2.3.22 $\models B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) = 0 \leftrightarrow B_{\pi a}\varphi = B_\pi\langle a\rangle\varphi$

PROOF Given a pointed model $\mathcal{M}, s$, we only need to show that $\mathcal{M}, s \models B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) = 0$ iff $\mathcal{M}, s \models B_{\pi a}\varphi = B_\pi\langle a\rangle\varphi$. This is obvious if $I^{\mathcal{M}}|^{\pi a} = \emptyset$. Next, we only focus on the case of $I^{\mathcal{M}}|^{\pi a} \neq \emptyset$. We have the following:

$$
\begin{array}{l}
\quad \mathcal{M}, s \not\models B_{\pi a}\varphi = B_\pi\langle a\rangle\varphi \\
\Leftrightarrow \mu^{\mathcal{M}}_{\pi a}([\![\varphi]\!]^{\mathcal{M}|^{\pi a}}) < \mu^{\mathcal{M}}_{\pi}([\![\langle a\rangle\varphi]\!]^{\mathcal{M}|^{\pi}}) \qquad \text{by Proposition 2.3.21} \\
\Leftrightarrow \text{there exists } s' \in I^{\mathcal{M}}|^{\pi} \text{ such that } \mathcal{M}|^{\pi a}, t' \models \varphi \text{ for some } t' \in R^{\mathcal{M}}_a(s) \text{ and} \\
\quad \Big(\sum\limits_{t\in([\![\varphi]\!]^{\mathcal{M}^{\pi a}}\cap R^{\mathcal{M}}_a(s))} Pr^{\mathcal{M}}(s,a,t)\Big) < 1 \quad \text{by the proof of Proposition 2.3.21} \\
\Leftrightarrow \text{there exists } s' \in I^{\mathcal{M}}|^{\pi} \text{ such that } \mathcal{M}|^{\pi a}, t' \models \varphi \text{ for some } t' \in R^{\mathcal{M}}_a(s) \text{ and} \\
\quad \mathcal{M}|^{\pi a}, t \not\models \varphi \text{ for some } t \in R^{\mathcal{M}}_a(s) \\
\Leftrightarrow \text{there exists } s' \in I^{\mathcal{M}}|^{\pi} \text{ such that } \mathcal{M}|^{\pi}, s \models \langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi \\
\Leftrightarrow \mu^{\mathcal{M}}_{\pi}([\![\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi]\!]^{\mathcal{M}|^{\pi}}) > 0 \\
\Leftrightarrow \mathcal{M}, s \not\models B_\pi(\langle a\rangle\varphi \wedge \langle a\rangle\neg\varphi) = 0
\end{array}
$$

□

The next axiom for which we prove soundness is ITSP. This axiom is a scheme for many different formulas. We abstract from the (in)equality expressed. We call the axiom introspection, because it is closely related to the usual axioms 4 and 5 in epistemic logic that express positive and negative introspection, respectively. Since an inequality is a negation, this scheme captures both positive and negative introspection in our probabilistic setting.

Proposition 2.3.23 $\models (\sum_{i=1}^n q_i B_{\pi\pi_i'}\varphi_i \bowtie qB_\pi\top) \rightarrow B_\pi(\sum_{i=1}^n q_i B_{\pi_i'}\varphi_i \bowtie q) = B_\pi\top$

PROOF If $\mathcal{M}, s \vDash \sum_{i=1}^n q_i B_{\pi\pi'_i}\varphi_i \bowtie qB_\pi\top$, namely, $\sum_{i=1}^n q_i \cdot \mu^{\mathcal{M}}_{\pi\pi'}([\![\varphi_i]\!]^{\mathcal{M}|^{\pi\pi'}}) \bowtie q \cdot \mu^{\mathcal{M}}_{\pi}([\![\top]\!]^{\mathcal{M}|^{\pi}})$, we need to show $\mathcal{M}, s \vDash B_\pi(\sum_{i=1}^n q_i B_{\pi'_i}\varphi_i \bowtie q) = B_\pi\top$, namely, $\mu^{\mathcal{M}}_{\pi}([\![\sum_{i=1}^n q_i B_{\pi'_i}\varphi_i \bowtie q]\!]^{\mathcal{M}|^{\pi}}) = \mu^{\mathcal{M}}_{\pi}([\![\top]\!]^{\mathcal{M}|^{\pi}})$. If $I^{\mathcal{M}}|^{\pi\pi'} = \emptyset$, it is obvious.

Next, we focus on the situation of $I^{\mathcal{M}}|^{\pi\pi'} \neq \emptyset$. To show $\mu^{\mathcal{M}}_{\pi}([\![\sum_{i=1}^n q_i B_{\pi'_i}\varphi_i \bowtie q]\!]^{\mathcal{M}|^{\pi}}) = \mu^{\mathcal{M}}_{\pi}([\![\top]\!]^{\mathcal{M}|^{\pi}})$, we only need to show that $[\![\sum_{i=1}^n q_i B_{\pi'_i}\varphi_i \bowtie q]\!]^{\mathcal{M}|^{\pi}} = I^{\mathcal{M}}|^{\pi}$. By semantics, we only need to show $\sum_{i=1}^n q_i \cdot \mu^{\mathcal{M}|^{\pi}}_{\pi'_i}([\![\varphi_i]\!]^{\mathcal{M}|^{\pi}|^{\pi'_i}}) \bowtie q$. Since $I^{\mathcal{M}}|^{\pi\pi'} \neq \emptyset$, it follows by Propositions 2.3.20 and 2.3.19 that $\mu^{\mathcal{M}|^{\pi}}_{\pi'}([\![\varphi_i]\!]^{\mathcal{M}|^{\pi}|^{\pi'}}) = \mu^{\mathcal{M}}_{\pi\pi'}([\![\varphi_i]\!]^{\mathcal{M}|^{\pi\pi'}})/\mu^{\mathcal{M}}_{\pi}(I^{\mathcal{M}}|^{\pi})$. Therefore, we have $\sum_{i=1}^n q_i \cdot \mu^{\mathcal{M}|^{\pi}}_{\pi'_i}([\![\varphi_i]\!]^{\mathcal{M}|^{\pi}|^{\pi'_i}}) \bowtie q$ if and only if $\sum_{i=1}^n q_i \cdot \mu^{\mathcal{M}}_{\pi\pi'}([\![\varphi_i]\!]^{\mathcal{M}|^{\pi\pi'}}) \bowtie q \cdot \mu^{\mathcal{M}}_{\pi}([\![\top]\!]^{\mathcal{M}|^{\pi}})$. □

The last axiom for which we prove soundness is CP, an axiom about conditional probability. It expresses the relation between prior and posterior probability in our setting.

Proposition 2.3.24 $\vDash \langle a\rangle\top \to ([a](\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \bowtie q) \leftrightarrow \sum_{i=1}^n q_i B_{a\pi_i}\varphi_i \bowtie qB_a\top)$

PROOF Given a pointed $\mathcal{M}, s \vDash$ and $s \xrightarrow{a} t$ for some $t \in S$, we need to show that $\mathcal{M}, s \vDash [a](\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \bowtie q) \leftrightarrow \sum_{i=1}^n q_i B_{a\pi_i}\varphi_i \bowtie qB_a\top$. Since $\mathcal{M}, s \vDash [a](\sum_{i=1}^n q_i B_{\pi_i}\varphi_i \bowtie q)$ if and only if $\mathcal{M}|^a, t \vDash \sum_{i=1}^n q_i B_{\pi_i}\varphi_i \bowtie q$. Thus, we only need to show $\mathcal{M}|^a, t \vDash \sum_{i=1}^n q_i B_{\pi_i}\varphi_i \bowtie q$ if and only if $\mathcal{M}, s \vDash \sum_{i=1}^n q_i B_{a\pi_i}\varphi_i \bowtie qB_a\top$. It is obvious if $I^{\mathcal{M}}|^{a\pi_i} = \emptyset$ for all $1 \leq i \leq n$. Next we only focus on the case of $I^{\mathcal{M}}|^{a\pi_i} \neq \emptyset$ for all $1 \leq i \leq n$.

$$\begin{array}{ll}
 & \mathcal{M}|^a, t \vDash \sum_{i=1}^n q_i B_{\pi_i}\varphi_i \bowtie q \\
\Leftrightarrow & \sum_{i=1}^n q_i \mu^{\mathcal{M}|^a}_{\pi_i}([\![\varphi_i]\!]^{\mathcal{M}|^a|^{\pi_i}}) \bowtie q \\
\Leftrightarrow & \sum_{i=1}^n q_i \mu^{\mathcal{M}|^a}_{\pi_i}([\![\varphi_i]\!]^{\mathcal{M}|^{a\pi_i}}) \bowtie q \quad \text{by Proposition 2.3.20} \\
\Leftrightarrow & \sum_{i=1}^n q_i/\mu^{\mathcal{M}}_{a}(I^{\mathcal{M}}|^a) \cdot \mu^{\mathcal{M}}_{a\pi_i}([\![\varphi_i]\!]^{\mathcal{M}|^{a\pi_i}}) \bowtie q \quad \text{by Proposition 2.3.19} \\
\Leftrightarrow & \sum_{i=1}^n q_i \mu^{\mathcal{M}}_{a\pi_i}([\![\varphi_i]\!]^{\mathcal{M}|^{a\pi_i}}) \bowtie q\mu^{\mathcal{M}}_{a}(I^{\mathcal{M}}|^a) \quad \text{due to } \mu^{\mathcal{M}}_{a}(I^{\mathcal{M}}|^a) > 0 \\
\Leftrightarrow & \mathcal{M}, s \vDash \sum_{i=1}^n q_i B_{a\pi_i}\varphi_i \bowtie qB_a\top
\end{array}$$

□

Now we are ready to prove the soundness lemma, which can be proven by induction on the length of the proof. We will leave the proof to the reader.

Theorem 2.3.25 (Soundness) *For each formula φ, $\vdash \varphi$ implies $\vDash \varphi$.*

2.4 Completeness

In this section we show that the axiomatization presented in Section 2.3.1 is complete with respect to the semantics we presented in Section 2.2. One important strategy that has been employed to prove completeness for dynamic epistemic logic is to use reduction axioms (see for instance van Benthem et al. (2006)). Reduction axioms are a way of relating what is the case after an announcement to what is the case before an announcement. Unfortunately that strategy is not available here because the agent cannot know beforehand what will be the case after an action without any information about the model. We will, however, try to reduce the language to certain forms and prove completeness with respect to this restricted language.

Our proof of the completeness for LCPP is divided into three steps: first, we reduce the language $\mathcal{L}_{\text{LCPP}}$ to its subset $\mathcal{L}^{\text{B}}_{\text{LCPP}}$; second, we define nonstandard models and show that each formula $\varphi \in \mathcal{L}^{\text{B}}_{\text{LCPP}}$ is satisfiable in nonstandard models implies it is also satisfiable in standard models defined in Section 2.2; third, we construct a canonical nonstandard model and show the truth lemma.

2.4.1 Normal form

In this subsection, we will show that each formula can be reduced to a certain form. Before defining the language of that form, we first define the language without probability, which is the normal language of modal logic.

Definition 2.4.1 *The language $\mathcal{L}^{\texttt{B-Free}}_{LCPP}$ is defined as the following BNF:*

$$\psi ::= p \mid \neg\psi \mid (\psi \wedge \psi) \mid [a]\psi$$

where $p \in \mathbf{P}$ and $a \in \mathbf{Act}$.

With the language $\mathcal{L}^{\texttt{B-Free}}_{\text{LCPP}}$, we now define the language with special form.

Definition 2.4.2 *The language $\mathcal{L}^{B}_{LCPP}$ is defined as the following BNF:*

$$\varphi ::= \psi \mid q_1 B_{\pi_1}\psi + \cdots + q_n B_{\pi_n}\psi \geq q \mid \neg\varphi \mid (\varphi \wedge \varphi)$$

where $\psi \in \mathcal{L}^{\texttt{B-Free}}_{LCPP}$, $\pi_i \in \mathbf{Act}^*$, and $q, q_i \in \mathbb{Q}$ for each $1 \leq i \leq n$.

It is obvious that $\mathcal{L}^{\texttt{B-Free}}_{\text{LCPP}} \subset \mathcal{L}^{\text{B}}_{\text{LCPP}}$ and $\mathcal{L}^{\text{B}}_{\text{LCPP}} \subset \mathcal{L}_{\text{LCPP}}$. Please note that for each formula in $\mathcal{L}^{\text{B}}_{\text{LCPP}}$ the probability operator B_π only appears outside a modal logic formula. In other words, there are no nesting occurrences between B_π and B_π, and between B_π and $[a]$.

The rest of this subsection will show that each formula $\varphi \in \mathcal{L}_{\text{LCPP}}$ can be reduced to be equivalent to a formula $\varphi' \in \mathcal{L}^{\text{B}}_{\text{LCPP}}$. To make the proof easier, we first define a notion of *conjunctive normal form* in a similar way as it is defined in propositional logic.

Definition 2.4.3 (Conjunctive Normal Form) *A formula $\varphi \in \mathcal{L}_{LCPP}$ is in conjunctive normal form if it is a conjunction of disjunctions of 'literals', where a 'literal' is a formula of the form p, $\neg p$, $[a]\psi$, $\neg[a]\psi$, $\sum_{i=1}^{n} q_i B_{\pi_i}\psi_i \geq q$ or $\neg(\sum_{i=1}^{n} q_i B_{\pi_i}\psi_i \geq q)$, where ψ and ψ_i are also in normal form.*

Since modal formulas and probability formulas are seen as propositional letters in conjunctive normal form, it is routine to show the following proposition.

Proposition 2.4.4 *For each formula $\varphi \in \mathcal{L}_{LCPP}$, there exists a formula $\varphi' \in \mathcal{L}_{LCPP}$ such that $\vdash \varphi \leftrightarrow \varphi'$ and φ' is in conjunctive normal form.*

Next, we define a notion which reflects a formula's complexity on the depth of nesting between probability operator and action operator.

Definition 2.4.5 (Nesting Degree) *Degree of a formula or an item is defined as follows.*

$$
\begin{array}{lcl}
d(p) & = & 0 \\
d(\neg\varphi) & = & d(\varphi) \\
d(\varphi \wedge \psi) & = & max\{d(\varphi), d(\psi)\} \\
d([a]\varphi) & = & \begin{cases} 1 + d(\varphi) & \text{if a probability term occurs in } \varphi \\ 0 & \text{else} \end{cases} \\
d(\sum_{i=1}^{n} q_i B_{\pi_i'} \varphi_i \geq q) & = & max\{d(B_{\pi_i'}\varphi_i) \mid 1 \leq i \leq n\} \\
d(B_\pi \varphi) & = & \begin{cases} 1 + d(\varphi) & \text{if a probability term occurs in } \varphi \\ 0 & \text{else} \end{cases}
\end{array}
$$

Please note that $d(\varphi) = 0$ for each $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$. Moreover, for each $\varphi \in \mathcal{L}_{\mathrm{LCPP}}$ if $d(\varphi) = 0$ then $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$. Therefore, our task is to reduce $\varphi \in \mathcal{L}_{\mathrm{LCPP}}$ to a formula φ' such that $d(\varphi') = 0$. To make the proof shorter, we need the following two propositions.

Proposition 2.4.6 *Given $[a]\psi \in \mathcal{L}_{LCPP}$ and $d([a]\psi) = 1$, there exists a formula φ such that $d(\varphi) = 0$ and $\vdash \varphi \leftrightarrow [a]\psi$.*

PROOF Since $d([a]\psi) = 1$, ψ cannot be of the form $[b]\chi$ or $\neg[b]\chi$. Without loss of generality, we assume ψ is in conjunctive normal form and $[a]\psi := [a](\psi_1 \vee \cdots \vee \psi_n \vee \psi')$ where $d([a]\psi') = 0$ and for all $1 \leq i \leq n$, $\psi_i := \sum_{j=1}^{i_n} q_{i_j} B_{\pi_{i_j}} \chi_{i_j} \geq q_i$ and for all $1 \leq j \leq i_n$, $d(B_{\pi_{i_j}} \chi_{i_j}) = 0$. By induction on n, we will show that there exists a formula φ with $d(\varphi) = 0$ such that $\vdash [a]\psi \leftrightarrow \varphi$.

If $n = 1$, $[a]\psi := [a](\sum_{j=1}^{m} q_j B_{\pi_j} \chi_j \geq q \vee \psi')$. Let $\varphi := (\sum_{j=1}^{m} q_j B_{a\pi_j} \chi_j \geq qB_a\top) \vee [a]\psi'$ then we have $d(\varphi) = 0$. It follows by Proposition 2.3.8 that $\vdash [a]\psi \leftrightarrow \varphi$. If $[a]\psi := [a](\psi_1 \vee \cdots \vee \psi_{n+1} \vee \psi')$ where $\psi_i := \sum_{j=1}^{i_n} q_{i_j} B_{\pi_{i_j}} \chi_{i_j} \geq q_i$ for each $1 \leq i \leq n+1$. Let $\varphi' := (\sum_{j=1}^{1_n} q_{1_j} B_{a\pi_{1_j}} \chi_{1_j} \geq q_1 B_a \top) \vee [a](\psi_2 \vee \cdots \vee \psi_{n+1} \vee \psi')$ then we have $d(\sum_{j=1}^{1_n} q_{1_j} B_{a\pi_{1_j}} \chi_{1_j} \geq q_1 B_a \top) = 0$. It follows by Proposition 2.3.8 that $\vdash [a]\psi \leftrightarrow \varphi'$. By induction on n, it follows that there exists a formula φ'' such that $d(\varphi'' = 0)$ and $\vdash \varphi'' \leftrightarrow [a](\psi_2 \vee \cdots \vee \psi_{n+1} \vee \psi')$. Let $\varphi := (\sum_{j=1}^{1_n} q_{1_j} B_{a\pi_{1_j}} \chi_{1_j} \geq q_1 B_a \top) \vee \varphi''$ then we have $d(\varphi) = 0$. It follows by Proposition 2.3.4 that $\vdash \varphi \leftrightarrow \varphi'$. Since $\vdash [a]\psi \leftrightarrow \varphi'$, it follows that $\vdash [a]\psi \leftrightarrow \varphi$. □

Proposition 2.4.7 *Given $\varphi \in \mathcal{L}_{LCPP}$, $\varphi := \sum_{i=1}^{n} q_i B_{\pi_i} \varphi_i \geq q$, and $d(\varphi) = 1$, there exists a formula φ' such that $d(\varphi') = 0$ and $\vdash \varphi \leftrightarrow \varphi'$.*

PROOF Without loss of generality, assume that each φ_i $(1 \leq i \leq n)$ is in conjunctive normal form. Since $d(\varphi) = 1$, it follows that at least one φ_i has a probability literal for some $1 \leq i \leq n$. Assume that φ_1 has a probability literal, namely $\varphi_1 := (\sum_{j=1}^{m} B_{\pi_j'} \chi_j' \bowtie q' \vee \chi'') \wedge \chi'''$. Then φ is of the form $q_1 B_{\pi_1}((\sum_{j=1}^{m} B_{\pi_j'} \chi_j' \bowtie q' \vee \chi'') \wedge \chi''') + \sum_{i=2}^{n} q_i B_{\pi_i} \varphi_i \geq q$ where $d(B_{\pi_j'} \chi_j') = 0$ for all $1 \leq j \leq m$. Let k be the number of occurrences of probability literals in $\varphi_1, \cdots, \varphi_n$. We prove it by induction on k.

If $k = 1$, it follows that $d(B_\pi(\chi'' \wedge \chi''')) = 0$ and $d(B_{\pi_i} \varphi_i) = 0$ for all $2 \leq i \leq n$. Let $\psi_1 := q_1 B_{\pi_1} \chi''' + \sum_{i=2}^{n} B_{\pi_i} \varphi_i \geq q$ and $\psi_2 := q_1 B_{\pi_1}(\chi'' \wedge \chi''') + \sum_{i=2}^{n} B_{\pi_i} \chi_i \geq q$. It follows that $d(\psi_1) = d(\psi_2) = 0$. Let $\varphi' := ((\sum_{j=1}^{m} B_{\pi\pi_j'} \chi_j' \bowtie q' B_\pi \top) \wedge \psi_1) \vee$

$(\neg(\sum_{j=1}^{m} B_{\pi\pi'_j}\chi'_j \bowtie q'B_\pi\top) \wedge \psi_2)$. Since $d(\sum_{j=1}^{m} B_{\pi\pi'_j}\chi'_j \bowtie q'B_\pi\top) = 0$, it follows that $d(\varphi') = 0$. It follows by Proposition 2.3.13 that $\vdash \varphi \leftrightarrow \varphi'$.

If $k = h+1$ and $h > 0$, Let $\psi_1 := q_1 B_{\pi_1}\chi''' + \sum_{i=2}^{n} B_{\pi_i}\varphi_i \geq q$ and $\psi_2 := q_1 B_{\pi_1}(\chi'' \wedge \chi''') + \sum_{i=2}^{n} B_{\pi_i}\chi_i \geq q$. It follows that $d(\psi_2) = 1$ and the number of occurrences of probability literals in $\chi'' \wedge \chi''', \varphi_2, \cdots, \varphi_n$ is h. It follows by IH that there exists a formula ψ'_2 such that $d(\psi'_2) = 0$ and $\vdash \psi_2 \leftrightarrow \psi'_2$. If $d(\psi_1) = 1$, it follows that the number of occurrences of probability literals in $\chi''', \varphi_2, \cdots, \varphi_n$ is less than or equal to h. It follows by IH that there exists a formula ψ'_1 such that $d(\psi'_1) = 0$ and $\vdash \psi_1 \leftrightarrow \psi'_1$. Let $\varphi'' := ((\sum_{j=1}^{m} B_{\pi\pi'_j}\chi'_j \bowtie q'B_\pi\top) \wedge \psi_1) \vee (\neg(\sum_{j=1}^{m} B_{\pi\pi'_j}\chi'_j \bowtie q'B_\pi\top) \wedge \psi_2)$ and $\varphi' := ((\sum_{j=1}^{m} B_{\pi\pi'_j}\chi'_j \bowtie q'B_\pi\top) \wedge \psi'_1) \vee (\neg(\sum_{j=1}^{m} B_{\pi\pi'_j}\chi'_j \bowtie q'B_\pi\top) \wedge \psi'_2)$. Since $d(\sum_{j=1}^{m} B_{\pi\pi'_j}\chi'_j \bowtie q'B_\pi\top) = 0$, we have $d(\varphi') = 0$. It follows by Proposition 2.3.13 that $\vdash \varphi \leftrightarrow \varphi''$. It follows by Proposition 2.3.4 that $\vdash \varphi' \leftrightarrow \varphi''$. Therefore, we have $\vdash \varphi \leftrightarrow \varphi'$. □

Now, we are ready to prove the reduction proposition.

Proposition 2.4.8 (Reduction) *For each formula $\varphi \in \mathcal{L}_{LCPP}$, there exists a formula $\varphi' \in \mathcal{L}^{B}_{LCPP}$ such that $\vdash \varphi \leftrightarrow \varphi'$.*

PROOF We only need to show that there exists φ' such that $d(\varphi') = 0$ and $\vdash \varphi \leftrightarrow \varphi'$. We prove this by induction on $d(\varphi)$. It is obvious if $d(\varphi) = 0$. If $d(\varphi) = n+1$, without loss of generality, we assume φ is in conjunctive normal form. It follows by Proposition 2.3.4 that we only need to show that for each literal ψ in φ there exists a formula ψ' such that $\vdash \psi \leftrightarrow \psi'$ and $d(\psi') = 0$. By IH, this is straightforward if $d(\psi) = n$. If $d(\psi) = n+1$, ψ is of the form $[a]\psi'$, $\neg[a]\psi'$, $\sum_{i=1}^{n} q_i B_{\pi_i}\psi'_i \geq q$ or $\neg\sum_{i=1}^{n} q_i B_{\pi_i}\psi'_i \geq q$. We only focus on the case of $[a]\psi'$ and $\sum_{i=1}^{n} q_i B_{\pi_i}\psi'_i \geq q$; the other cases are similar.

If $\psi := [a]\psi'$ and $d([a]\psi') = n+1$, it follows that $d(\psi') = n$. By IH, it follows that there exists a formula χ' such that $\vdash \psi' \leftrightarrow \chi'$ and $d(\chi') = 0$. Thus, we have $\vdash \psi \leftrightarrow [a]\chi'$ and $d([a]\chi') \leq 1$. If $d([a]\chi') = 1$, it follows by Proposition 2.4.6 that there exists a formula χ such that $\vdash \chi \leftrightarrow [a]\chi'$ and $d(\chi) = 0$. It follows that $\vdash \psi \leftrightarrow \chi$.

If $\psi := \sum_{i=1}^{n} q_i B_{\pi_i}\psi'_i \geq q$ and $d(\psi) = n+1$, it follows that $d(\psi'_i) \leq n$ for all $1 \leq i \leq n$. By IH, it follows that for each ψ'_i there exists a formula χ'_i such that $\vdash \psi'_i \leftrightarrow \chi'_i$ and $d(\chi'_i) = 0$. It follows that $\vdash \psi \leftrightarrow \sum_{i=1}^{n} q_i B_{\pi_i}\chi'_i \geq q$ and $d(\sum_{i=1}^{n} q_i B_{\pi_i}\chi'_i \geq q) \leq 1$. If $d(\sum_{i=1}^{n} q_i B_{\pi_i}\chi'_i \geq q) = 1$, it follows by Proposition 2.4.7 that there exists a formula χ such that $\vdash \sum_{i=1}^{n} q_i B_{\pi_i}\chi'_i \geq q \leftrightarrow \chi$ and $d(\chi) = 0$. It follows that $\vdash \psi \leftrightarrow \chi$. □

Due to the soundness of $\mathbb{SLCPP}$, the above proposition also indicates that the expressive power of the full language $\mathcal{L}_{\text{LCPP}}$ is the same as its fragment $\mathcal{L}^{\text{B}}_{\text{LCPP}}$. The rest of this chapter will focus on formulas in $\mathcal{L}^{\text{B}}_{\text{LCPP}}$.

2.4.2 Nonstandard model

Recall the models and the semantics defined in Section 2.2. There are two kinds of probabilities in a model, and the probability representing the agent's initial uncertainty needs to be updated in the semantics. These will cause too much troubles if we directly construct the canonical model. Our strategy is working on alternative models, which are defined below. To avoid confusion, we call the model and the semantics defined in Section 2.2 standard model and standard semantics. In this subsection, firstly we will define nonstandard models and nonstandard semantics, then we will show that if a formula $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$ is satisfiable in nonstandard models, then it is also satisfiable in standard models.

Definition 2.4.9 (Nonstandard Model) *A nonstandard model, denoted as $\mathfrak{M}$, is a tuple $\mathfrak{M} = \langle S^{\mathfrak{M}}, R^{\mathfrak{M}}, I^{\mathfrak{M}}, \{\mu^{\mathfrak{M}}_{\pi} \mid \pi \in \mathbf{Act}^*\}, V^{\mathfrak{M}}\rangle$ such that*

- $S^{\mathfrak{M}} \neq \emptyset$ *is a finite set of states,*
- $R^{\mathfrak{M}} \subseteq S^{\mathfrak{M}} \times \mathbf{Act} \times S^{\mathfrak{M}}$,
- $I^{\mathfrak{M}}$ *is a non-empty subset of* $S^{\mathfrak{M}}$,
- $\mu^{\mathfrak{M}}_{\pi} : I^{\mathfrak{M}}|^{\pi} \to [0,1]$ *is a function such that*
  1. $\mu^{\mathfrak{M}}_{\epsilon}(I^{\mathfrak{M}}) = 1$ *and* $\mu^{\mathfrak{M}}_{\epsilon}(s) > 0$ *for each* $s \in I^{\mathfrak{M}}$*;*
  2. $\mu^{\mathfrak{M}}_{\pi a}(I^{\mathfrak{M}}|^{\pi a}) = \mu^{\mathfrak{M}}_{\pi}(\{s \in I^{\mathfrak{M}}|^{\pi} \mid R^{\mathfrak{M}}_a(s) \neq \emptyset\})$ *and* $\mu^{\mathfrak{M}}_{\pi a}(s) > 0$ *for each* $s \in I^{\mathfrak{M}}|^{\pi a}$*;*
  3. $\mu^{\mathfrak{M}}_{\pi a}(E) \leq \mu^{\mathfrak{M}}_{\pi}(\{s \in I^{\mathfrak{M}}|^{\pi} \mid \exists t \in E : s \xrightarrow{a} t\})$ *for each* $E \subseteq I^{\mathfrak{M}}|^{\pi a}$*;*
  4. $\mu^{\mathfrak{M}}_{\pi a}(E) < \mu^{\mathfrak{M}}_{\pi}(\{s \in I^{\mathfrak{M}}|^{\pi} \mid \exists t \in E : s \xrightarrow{a} t\})$ *for each* $E \subseteq I^{\mathfrak{M}}|^{\pi a}$ *such that* $R^{\mathfrak{M}}_a(s) \cap E \neq \emptyset$ *and* $R^{\mathfrak{M}}_a(s) \setminus E \neq \emptyset$ *for some* $s \in I^{\mathfrak{M}}|^{\pi}$*;*
- $V^{\mathfrak{M}} : \mathbf{P} \to \mathcal{P}(S^{\mathfrak{M}})$.

Please note that $I^{\mathfrak{M}}|^{\pi}$ is defined in the same way as it is in Section 2.2. There are two differences between nonstandard models and standard models: first, there are no transition probabilities in nonstandard models; second, the functions μ_{π} of standard models are calculated by two kinds of probabilities (see Definition 2.2.4) while they are predefined in nonstandard models, but intuitively they are the same thing. The requirements of the functions $\mu^{\mathfrak{M}}_{\pi}$ of nonstandard models make sure that there are indeed transition probabilities such that $\mu^{\mathfrak{M}}_{\pi}$ are calculated in the way shown in Definition 2.2.4.

Definition 2.4.10 (Nonstandard Semantics) *Let $\mathfrak{M}$ be a nonstandard model. Given a state $s \in S^{\mathfrak{M}}$ and a formula $\psi \in \mathcal{L}^{B\text{-}Free}_{LCPP}$, ψ being true at $\mathfrak{M}, s$, denoted as $\mathfrak{M}, s, \Vdash \psi$, is defined the same as the standard semantics. Given $s \in I^{\mathfrak{M}}$ and $\varphi \in \mathcal{L}^{B}_{LCPP}$, $\mathfrak{M}, s, \Vdash \varphi$ is defined as follows:*

$$\begin{array}{rcl} \mathfrak{M}, s \Vdash \neg\varphi & \iff & \mathfrak{M}, s \nVdash \varphi \\ \mathfrak{M}, s \Vdash (\varphi \wedge \psi) & \iff & \mathfrak{M}, s \Vdash \varphi \text{ and } \mathfrak{M}, s \Vdash \psi \\ \mathfrak{M}, s \Vdash \sum_{i=1}^{n} q_i B_{\pi_i}\varphi_i \geq q & \iff & \sum_{i=1}^{n} q_i \mu^{\mathfrak{M}}_{\pi_i}([\![\varphi_i]\!]^{\mathfrak{M}}_{\pi_i}) \geq q \end{array}$$

where $[\![\varphi]\!]^{\mathfrak{M}}_{\pi_i} = \{s \in I^{\mathfrak{M}}|^{\pi_i} \mid \mathfrak{M}, s \Vdash \varphi\}$.

Note that a probability formula is only evaluated on states in $I^{\mathfrak{M}}$. Moreover, there is no updating in nonstandard semantics. The reason is that in nonstandard semantics we only care about formulas in $\mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$. Formulas in $\mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$ cannot express the agent's uncertainty after she performs an action. Therefore, there is no need to update the model.

The following proposition is crucial, which shows that our strategy of working on nonstandard models does make sense.

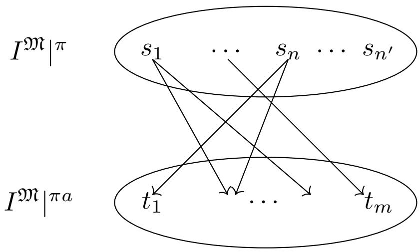
Proposition 2.4.11 *Given $\varphi \in \mathcal{L}^B_{LCPP}$, if it is satisfiable in nonstandard models then it is also satisfiable in standard models.*

PROOF Let $\mathfrak{M}$ be a finite nonstandard model such that $\mathfrak{M}, u \Vdash \varphi$ where $u \in I^{\mathfrak{M}}$. We need to show that there exists a standard pointed model in which φ is true. We only need to show the following claims.

The only differences between standard models and nonstandard models are probabilities. Recall Definition 2.2.4, and we know that in standard models, μ_a is calculated by the probability B and the probability Pr_a. Since μ_ϵ in nonstandard models is the same as B in standard models, the first claim is to show that there exist such functions Pr_a that Pr_a is a probability distribution and that μ_a in nonstandard models coincides with μ_a which is calculated by this Pr_a and μ_ϵ. The idea of the claim proof is that we list a set of inequalities based on the probability μ_a in nonstandard models and the conditions that Pr_a needs to satisfy, and then we show the inequality set is satisfiable by using Theorem A.2.7.

Claim 2.4.11.1 *Given $\pi a \in \mathbf{Act}^*$, if $I^{\mathfrak{M}}|^{\pi a} \neq \emptyset$, there exists a function $Pr^{\mathfrak{M}}_{\pi a} : R^{\mathfrak{M}}_a|_{(I^{\mathfrak{M}}|^{\pi} \times I^{\mathfrak{M}}|^{\pi a})} \to \mathbb{Q}^+$ such that $\sum_{t \in R^{\mathfrak{M}}_a(s)} Pr^{\mathfrak{M}}_{\pi a}(s,t) = 1$ for each $s \in I^{\mathfrak{M}}|^{\pi}$ where a is executable at s, and that $\sum_{\{s \in I^{\mathfrak{M}}|^{\pi} | t \in R^{\mathfrak{M}}_a(s)\}} \mu^{\mathfrak{M}}_{\pi}(s) \cdot Pr^{\mathfrak{M}}_{\pi a}(s,t) = \mu^{\mathfrak{M}}_{\pi a}(t)$ for each $t \in I^{\mathfrak{M}}|^{\pi a}$.*

Proof of claim 2.4.11.1: Intuitively, $Pr^{\mathfrak{M}}_{\pi a}(s,t)$ represents the probability of reaching t by performing a in s. Let $I^{\mathfrak{M}}|^{\pi} = \{s_1, \cdots, s_n, \cdots, s_{n'}\}$ such that a is executable at each s_i where $1 \leq i \leq n$ and a is unexecutable at each s_j where $n < j \leq n'$. Let $I^{\mathfrak{M}}|^{\pi a} = \{t_1, \cdots, t_m\}$ then the relation $R^{\mathfrak{M}}_a$ on $I^{\mathfrak{M}}|^{\pi} \times I^{\mathfrak{M}}|^{\pi a}$ can be roughly depicted as follows.



We now describe a set of linear inequalities over variables of the form $x_{(i,j)}$ for $(s_i, t_j) \in R^{\mathfrak{M}}_a|_{(I^{\mathfrak{M}}|^{\pi} \times I^{\mathfrak{M}}|^{\pi a})}$. The variable $x_{(i,j)}$ represents $\mu^{\mathfrak{M}}_{\pi}(s_i) \cdot Pr^{\mathfrak{M}}_{\pi a}(s_i, t_j)$. For each $(s_i, t_j) \in R^{\mathfrak{M}}_a|_{(I^{\mathfrak{M}}|^{\pi} \times I^{\mathfrak{M}}|^{\pi a})}$, to make sure $Pr^{\mathfrak{M}}_{\pi a}(s_i, t_j) > 0$, we only need to request that

$$x_{(i,j)} > 0. \tag{2.1}$$

For each $1 \leq i \leq n$, to make sure $\sum_{t_j \in R^{\mathfrak{M}}_a(s_i)} Pr^{\mathfrak{M}}_{\pi a}(s_i, t_j) = 1$, we only need to request that

$$\sum_{t_j \in R^{\mathfrak{M}}_a(s_i)} x_{(i,j)} = \mu^{\mathfrak{M}}_{\pi}(s_i). \tag{2.2}$$

For each $1 \leq j \leq m$, to make sure $\sum_{\{s_i \in I^{\mathfrak{M}}|^{\pi} | t_j \in R^{\mathfrak{M}}_a(s_i)\}} \mu^{\mathfrak{M}}_{\pi}(s_i) \cdot Pr^{\mathfrak{M}}_{\pi a}(s_i, t_j) =$

$\mu^{\mathfrak{M}}_{\pi a}(t_j)$, we only need to request that

$$\sum_{\{s_i \in I^{\mathfrak{M}}|^{\pi} | t_j \in R^{\mathfrak{M}}_a(s_i)\}} x_{(i,j)} = \mu^{\mathfrak{M}}_{\pi a}(t_j). \tag{2.3}$$

Next, we only need to show that the set S of linear inequalities described in (2.1) to (2.3) has a solution. By Theorem A.2.7, we only need to show that

$$\sum_{(s_i,t_j)\in R^{\mathfrak{M}}_a|_{\mathcal{U}^{\mathfrak{M}}|^{\pi} \times \mathcal{U}^{\mathfrak{M}}|^{\pi a}}} 0x_{(i,j)} > 0 \tag{2.4}$$

is not a possible legal linear combination of S. If possible, let $\mu^{\mathfrak{M}}_{\pi}(s_i) = a_i$ and $\mu^{\mathfrak{M}}_{\pi a}(t_j) = b_j$ then there exists a scheme (cf. Definition A.2.6) of S as shown in Table 2.2 such that

$$u_{(i',j')} > 0 \text{ for some } (s_{i'}, t_{j'}) \in R^{\mathfrak{M}}_a|_{\mathcal{U}^{\mathfrak{M}}|^{\pi} \times \mathcal{U}^{\mathfrak{M}}|^{\pi a}}; \tag{2.5}$$

$$d_{(i,j)} = u_{(i,j)} + r_i + w_j = 0 \text{ for each } (s_i, t_j) \in R^{\mathfrak{M}}_a|_{\mathcal{U}^{\mathfrak{M}}|^{\pi} \times \mathcal{U}^{\mathfrak{M}}|^{\pi a}}; \tag{2.6}$$

$$d = -u_0 + r_1 a_1 + \cdots + r_n a_n + w_1 b_1 + \cdots + w_m b_m = 0 \tag{2.7}$$

where $r_i = r'_{2i-1} - r'_{2i}$ and $w_j = w'_{2j-1} - w'_{2j}$ for all $1 \le i \le n$ and $1 \le j \le m$.

Let a group G be a minimal subset of $I^{\mathfrak{M}}|^{\pi a}$ such that for each $t \in G$ if $s \xrightarrow{a} t$ and $s \xrightarrow{a} t'$ for some $s \in I^{\mathfrak{M}}|^{\pi}$ then $t' \in G$. It is obvious that $I^{\mathfrak{M}}|^{\pi a}$ can be divided into several such groups. Without loss of generality, assume one of these group is $\{t_1, \cdots, t_h\}$, and we will write it as $\{1, \cdots, h\}$ for abbreviation. For each $j \in G$, let $D_j = \{1 \le i \le n \mid (s_i, t_j) \in R^{\mathfrak{M}}_a|_{\mathcal{U}^{\mathfrak{M}}|^{\pi} \times \mathcal{U}^{\mathfrak{M}}|^{\pi a}}\}$, which is the set of all the numbers i such that $s_i \in I^{\mathfrak{M}}|^{\pi}$ and $s_i \xrightarrow{a} t_j$. For each $i \in D_j$, since $d_{(i,j)} = 0$ and $u_{(i,j)} \ge 0$, it follows that $w_j \le -r_i$. Given $j \in G$, we use r_{w_j} to denote the maximal number in $\{r_i \mid i \in D_j\}$. Since $w_j \le -r_i$ for all $i \in D_j$, it follows that $w_j \le -r_{w_j}$. Without loss of generality, we assume that $r_{w_1} \le \cdots \le r_{w_h}$. We use $D_{1,j}$ as an abbreviation for $D_1 \cup \cdots \cup D_j$. It follows by Definition 2.4.9 that $\sum_{i\in D_{1,h}} a_i = b_1 + \cdots + b_h$ and that $\sum_{i\in D_{1,k}} a_i > b_1 + \cdots + b_k$ for each $k < h$. We then have the following:

$$\begin{aligned}
&\textstyle\sum_{i\in D_{1,h}} r_i a_i + \sum_{j=1}^{h} w_j b_j \\
&\le \textstyle\sum_{i\in D_{1,h}} r_i a_i + \sum_{j=1}^{h} -r_{w_j} b_j \\
&\le r_{w_1}(\textstyle\sum_{i\in D_1} a_i - b_1) + \sum_{i\in D_{1,h}\setminus D_1} r_i a_i + \sum_{j=2}^{h} -r_{w_j} b_j \\
&\le r_{w_2}(\textstyle\sum_{i\in D_{1,2}} a_i - b_1 - b_2) + \sum_{i\in D_{1,h}\setminus D_{1,2}} r_i a_i + \sum_{j=3}^{m} -r_{w_j} b_j \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\le r_{w_h} \cdot (\textstyle\sum_{i\in D_{1,h}} a_i + \sum_{j=1}^{h} -b_j) \\
&= 0
\end{aligned} \tag{2.8}$$

For the reason why for each $1 \le k < h$,
$r_{w_k}(\sum_{i\in D_{1,k}} a_i + \sum_{j=1}^{k} -b_j) + \sum_{i\in D1,h\setminus D_{1,k}} r_i a_i + \sum_{j=k+1}^{h} -r_{w_j} b_j$
$\le r_{w_{k+1}}(\sum_{i\in D_{1,(k+1)}} a_i + \sum_{j=1}^{k+1} -b_j) + \sum_{i\in D_{1,h}\setminus D_{1,(k+1)}} r_i a_i + \sum_{j=k+2}^{h} -r_{w_j} b_j$

$$
\begin{array}{rcll}
u_0 \geq 0: & \sum_{(s_i,t_j)\in R_a^{\mathfrak{M}}|_{\mathcal{U}^{\mathfrak{M}}|^{\pi}\times\mathcal{U}^{\mathfrak{M}}|^{\pi a}}} 0x_{(i,j)} & > -1 & (0,0)\\
u_{(1,1)} \geq 0: & x_{(1,1)} & > 0 & (1,1)\\
\cdots & \cdots & \cdots & \cdots\\
u_{(i,j)} \geq 0: & x_{(i,j)} & > 0 & (i,j)\\
r'_1 \geq 0: & \sum_{t_j\in R_a^{\mathfrak{M}}(s_1)} x_{(1,j)} & \geq a_1 & (1,+)\\
r'_2 \geq 0: & \sum_{t_j\in R_a^{\mathfrak{M}}(s_1)} -x_{(1,j)} & \geq -a_1 & (1,-)\\
\cdots & \cdots & \cdots & \cdots\\
r'_{2n-1} \geq 0: & \sum_{t_j\in R_a^{\mathfrak{M}}(s_n)} x_{(n,j)} & \geq a_n & (n,+)\\
r'_{2n} \geq 0: & \sum_{t_j\in R_a^{\mathfrak{M}}(s_n)} -x_{(n,j)} & \geq -a_n & (n,-)\\
w'_1 \geq 0: & \sum_{\{s_i\in I^{\mathfrak{M}}|^{\pi}|t_1\in R_a^{\mathfrak{M}}(s_i)\}} x_{(i,1)} & \geq b_1 & (+,1)\\
w'_2 \geq 0: & \sum_{\{s_i\in I^{\mathfrak{M}}|^{\pi}|t_1\in R_a^{\mathfrak{M}}(s_i)\}} -x_{(i,1)} & \geq -b_1 & (-,1)\\
\cdots & \cdots & \cdots & \cdots\\
w'_{2m-1} \geq 0: & \sum_{\{s_i\in I^{\mathfrak{M}}|^{\pi}|t_m\in R_a^{\mathfrak{M}}(s_i)\}} x_{(i,m)} & \geq b_m & (+,m)\\
w'_2m \geq 0: & \sum_{\{s_i\in I^{\mathfrak{M}}|^{\pi}|t_m\in R_a^{\mathfrak{M}}(s_i)\}} -x_{(i,m)} & \geq -b_m & (-,m)\\
\hline
 & d_{(1,1)}x_{(1,1)} + \cdots + d_{(i,j)}x_{(i,j)} & > d & (0).
\end{array}
$$

Table 2.2: Scheme

(Scheme is a systematical way to get a logical consequence of a set of inequalities (cf. A.2.6). In this table, the inequality (0) is a logical consequence of all the inequalities above the line provided that some $u_{(i',j')}$ is positive. Each of $(0,0)$ ... $(-,m)$ stands for an inequality. For example, $(1,1)$ stands for $u_{(1,1)} \cdot x_{(1,1)} > u_{(1,1)} \cdot 0$ where $u_{(1,1)}$ is a non-negative coefficient. The inequality $(0,0)$ is based on $0 > -1$; each of $(1,1)$ to (i,j) is based on the inequality (2.1); the inequalities $(1,+)$ and $(1,-)$ are based on the equation (2.2); the inequalities $(+,1)$ and $(-,1)$ are based on the equation (2.3).)

we have the following.

$$
\begin{aligned}
& r_{w_k}(\textstyle\sum_{i\in D_{1,k}} a_i + \sum_{j=1}^{k} -b_j) + \sum_{i\in D1,h\setminus D_{1,k}} r_i a_i + \sum_{j=k+1}^{h} -r_{w_j} b_j \\
& \leq r_{w_k}(\textstyle\sum_{i\in D_{1,k}} a_i + \sum_{j=1}^{k} -b_j) + r_{w_{k+1}} \sum_{i\in D_{k+1}\setminus D_{1,k}} a_i \\
& \qquad\qquad + \textstyle\sum_{i\in D_{1,h}\setminus D_{1,(k+1)}} r_i a_i + \sum_{j=k+1}^{h} -r_{w_j} b_j \\
& \leq r_{w_{k+1}}(\textstyle\sum_{i\in D_{1,k}} a_i + \sum_{j=1}^{k} -b_j) + r_{w_{k+1}} \sum_{i\in D_{k+1}\setminus D_{1,k}} a_i \\
& \qquad\qquad + \textstyle\sum_{i\in D_{1,h}\setminus D_{1,(k+1)}} r_i a_i + \sum_{j=k+1}^{m} -r_{w_j} b_j \\
& = r_{w_{k+1}}(\textstyle\sum_{i\in D_{1,(k+1)}} a_i + \sum_{j=1}^{k+1} -b_j) + \sum_{i\in D_{1,h}\setminus D_{1,(k+1)}} r_i a_i \\
& \qquad\qquad + \textstyle\sum_{j=k+2}^{h} -r_{w_j} b_j
\end{aligned}
\tag{2.9}
$$

Because of the property of group, it follows that if G and G' are two different groups, and $t \in G$, $t' \in G'$ then $D_t \cap D_{t'} = \emptyset$. Assuming $I^{\mathfrak{M}}|^{\pi a}$ is divided into l groups, it follows that

$$d = -u_0 + \sum_{1\leq k\leq l} (\sum_{i\in D_{G_k}} r_i a_i + \sum_{j\in G_k} w_j b_j) \tag{2.10}$$

Since $d = 0$, $-u_0 \leq 0$ and $\sum_{i\in D_G} r_i a_i + \sum_{j\in G} w_j b_j \leq 0$ for each group G, it follows that $u_0 = 0$ and $\sum_{i\in D_G} r_i a_i + \sum_{j\in G} w_j b_j = 0$. It follows that each inequality of (2.8) or (2.9) equals 0, especially,

$$\sum_{i\in D_G} r_i a_i + \sum_{j\in G} w_j b_j = \sum_{i\in D_G} r_i a_i + \sum_{j\in G} -r_{w_j} b_j = 0. \tag{2.11}$$

Moreover, by (2.9), we have that for each $1 \leq k < h$,

$$r_{w_k}(\sum_{i\in D_{1,k}} a_i + \sum_{j=1}^{k} -b_j) = r_{w_{k+1}}(\sum_{i\in D_{1,k}} a_i + \sum_{j=1}^{k} -b_j), \tag{2.12}$$

$$\sum_{i\in D_{k+1}\setminus D_{1,k}} r_i a_i = r_{w_{k+1}} \sum_{i\in D_{k+1}\setminus D_{1,k}} a_i. \tag{2.13}$$

Since $\sum_{i\in D_{1,k}} a_i + \sum_{j=1}^{k} -b_j > 0$, it follows by (2.12) that $r_{w_k} = r_{w_{k+1}}$ for each $1 \leq k < h$. Next, we will show that for each $1 \leq k \leq h$, namely $t_k \in G$, $i \in D_k$ implies $r_i = r_{w_k}$. For the case of $k = 1$, it is obvious from (2.8). For the case of $k+1$, if $i \in D_{k+1} \setminus D_{1,k}$, it is obvious from (2.13). If $i \notin D_{k+1} \setminus D_{1,k}$, it follow by IH that $r_i = r_{w_{k'}}$ for some $k' \leq k$. Since $r_{w_k} = r_{w_{k+1}}$ for all $1 \leq k < h$, it follows that r, it follows that $r_i = r_{w_{k+1}}$.

By (2.5), we have known that $u_{(i',j')} > 0$. Since $r_i = r_{i'}$ and $d_{(i,j')} = u_{(i,j')} + r_i + w_{j'} = 0$ for all $i \in D_{j'}$, it follows that $u_{(i,j')} = u_{i',j'}$ for all $i \in D_{j'}$. Since $u_{(i',j')} > 0$, it follows that $r_{w_{j'}} + w_{j'} < 0$, namely $w_{j'} < -r_{w_{j'}}$. Thus, for the group G such that $j' \in G$, we have the following

$$\sum_{i\in D_G} r_i a_i + \sum_{j\in G} w_j b_j < \sum_{i\in D_G} r_i a_i + \sum_{j\in G} -r_{w_j} b_j.$$

This is in contradiction with (2.11). Therefore, (2.4) cannot be a legal linear combination of S, and it follows by Theorem A.2.7 that S has a solution.

Therefore, there exists a function $Pr^{\mathfrak{M}}_{\pi a}$ on $R^{\mathfrak{M}}_a|_{(I^{\mathfrak{M}}|^{\pi}\times I^{\mathfrak{M}}|^{\pi a})}$ which is defined as $Pr^{\mathfrak{M}}_{\pi a}(s_i,t_j) = x_{(i,j)}/\mu^{\mathfrak{M}}_{\pi}(s_i)$ for each $(s_i,t_j) \in R^{\mathfrak{M}}_a|_{(I^{\mathfrak{M}}|^{\pi}\times I^{\mathfrak{M}}|^{\pi a})}$. It follows from (2.1) to (2.3) that $Pr^{\mathfrak{M}}_{\pi a} : R^{\mathfrak{M}}_a|_{(I^{\mathfrak{M}}|^{\pi}\times I^{\mathfrak{M}}|^{\pi a})} \to \mathbb{Q}^+$ such that $\sum_{t\in R^{\mathfrak{M}}_a(s)} Pr^{\mathfrak{M}}_{\pi a}(s,t) = 1$ for each $s \in I^{\mathfrak{M}}|^{\pi}$ where a is executable at s, and that $\sum_{\{s\in I^{\mathfrak{M}}|^{\pi} | t\in R^{\mathfrak{M}}_a(s)\}} \mu^{\mathfrak{M}}_{\pi}(s)\cdot Pr^{\mathfrak{M}}_{\pi a}(s,t) = \mu^{\mathfrak{M}}_{\pi a}(t)$ for each $t \in I^{\mathfrak{M}}|^{\pi a}$. ■

Please recall that an execution path $\sigma \in EP_{\mathfrak{M}}(a_1\cdots a_n)$ is an alternating sequence of states and actions, $s_0a_1\cdots s_n$, where $s_0 \in I^{\mathfrak{M}}$ and $s_{i-1} \xrightarrow{a_i} s_i$ for each $1 \le i \le n$ (see Definition 2.2.3). Given $\sigma := s_0a_1\cdots s_n$, we use $T(\sigma)$ to denote the last state s_n and $\rho(\sigma)$ to the action sequence $a_1\cdots a_n$. Given $t \in I^{\mathfrak{M}}|^{\pi}$, let $[\sigma]^{\pi}_t$ denot the set $\{\sigma \in EP_{\mathfrak{M}}(\pi) \mid T(\sigma) = t\}$, which is the set of π-executions leading to t.

Next, we construct a standard model based on the execution paths of $\mathfrak{M}$. The standard model $\mathfrak{M}^\bullet$ is defined as follows.

$$\begin{array}{lcl}
S^{\mathfrak{M}^\bullet} & = & \{\sigma \in EP_{\mathfrak{M}}(\pi) \mid \pi \in \mathbf{Act}^*\} \\
R^{\mathfrak{M}^\bullet} & = & \{(\sigma,a,\sigma') \mid a \in \mathbf{Act}|_{\varphi}, \sigma' = \sigma a t\} \\
Pr^{\mathfrak{M}^\bullet}(\sigma,a,\sigma a t) & = & Pr^{\mathfrak{M}}_{\rho(\sigma)a}(T(\sigma),t) \\
I^{\mathfrak{M}^\bullet} & = & I^{\mathfrak{M}} \\
B^{\mathfrak{M}^\bullet} & = & \mu^{\mathfrak{M}}_{\epsilon} \\
V^{\mathfrak{M}^\bullet}(p) & = & \{\sigma \mid T(\sigma) \in V^{\mathfrak{M}}(p)\}
\end{array}$$

where $Pr^{\mathfrak{M}}_{\rho(\sigma)a}$ is the function which is shown in Claim 2.4.11.1.

Claim 2.4.11.2 *For each $\pi \in \mathbf{Act}^*$ and each $t \in I^{\mathfrak{M}}|^{\pi}$, we have $\mu^{\mathfrak{M}^\bullet}_{\pi}([\sigma]^{\pi}_t) = \mu^{\mathfrak{M}}_{\pi}(t)$.*

Proof of claim 2.4.11.2: By the definition of $\mathfrak{M}^\bullet$, it is obvious that $\sigma \in I^{\mathfrak{M}^\bullet}|^{\pi}$ iff $T(\sigma) \in I^{\mathfrak{M}}|^{\pi}$ for each $\sigma \in S^{\mathfrak{M}^\bullet}$. By induction on π we will show that $\mu^{\mathfrak{M}^\bullet}_{\pi}([\sigma]^{\pi}_t) = \mu^{\mathfrak{M}}_{\pi}(t)$. It is obvious for the case of ϵ. For the case of πa, we have the following.

$$\begin{aligned}
\mu^{\mathfrak{M}^\bullet}_{\pi a}([\sigma a t]^{\pi a}_t) &= \mu^{\mathfrak{M}^\bullet}_{\pi a}(\{\sigma' a t \mid s \in I^{\mathfrak{M}}|^{\pi}, t \in R^{\mathfrak{M}}_a(s), \sigma' \in [\sigma]^{\pi}_s\}) \\
&= \sum_{\{s\in I^{\mathfrak{M}}|^{\pi} | t\in R^{\mathfrak{M}}_a(s)\}} \sum_{\sigma'\in[\sigma]^{\pi}_s} \mu^{\mathfrak{M}^\bullet}_{\pi a}(\sigma' a t) \\
&= \sum_{\{s\in I^{\mathfrak{M}}|^{\pi} | t\in R^{\mathfrak{M}}_a(s)\}} \sum_{\sigma'\in[\sigma]^{\pi}_s} \mu^{\mathfrak{M}^\bullet}_{\pi}(\sigma')\cdot Pr^{\mathfrak{M}^\bullet}(\sigma',a,\sigma' a t) \\
&= \sum_{\{s\in I^{\mathfrak{M}}|^{\pi} | t\in R^{\mathfrak{M}}_a(s)\}} \sum_{\sigma'\in[\sigma]^{\pi}_s} \mu^{\mathfrak{M}^\bullet}_{\pi}(\sigma')\cdot Pr^{\mathfrak{M}}_{\pi a}(s,t) \\
&= \sum_{\{s\in I^{\mathfrak{M}}|^{\pi} | t\in R^{\mathfrak{M}}_a(s)\}} Pr^{\mathfrak{M}}_{\pi a}(s,t) \sum_{\sigma'\in[\sigma]^{\pi}_s} \mu^{\mathfrak{M}^\bullet}_{\pi}(\sigma') \\
&= \sum_{\{s\in I^{\mathfrak{M}}|^{\pi} | t\in R^{\mathfrak{M}}_a(s)\}} Pr^{\mathfrak{M}}_{\pi a}(s,t)\cdot \mu^{\mathfrak{M}}_{\pi}(s) && \text{(by IH)} \\
&= \mu^{\mathfrak{M}}_{\pi a}(t) && \text{(by Claim 2.4.11.1)}
\end{aligned}$$

Therefore, we have shown that $\mu^{\mathfrak{M}^\bullet}_{\pi}([\sigma]^{\pi}_t) = \mu^{\mathfrak{M}}_{\pi}(t)$ for each $t \in I^{\mathfrak{M}}|^{\pi}$. ■

Claim 2.4.11.3 *Given $\psi \in \mathcal{L}_{LCPP}^{\texttt{B-Free}}$, $\sigma \in S^{\mathfrak{M}^\bullet}$, and $\pi \in \mathbf{Act}^*$, if $\sigma \in I^{\mathfrak{M}^\bullet}|^\pi$, we have $\mathfrak{M}^\bullet|^\pi, \sigma \vDash \psi$ iff $\mathfrak{M}, T(\sigma) \Vdash \psi$.*

Proof of claim 2.4.11.3: It is easy to show by induction on ψ. ■

Claim 2.4.11.4 *For each formula $\psi \in \mathcal{L}_{LCPP}^{B}$ and each $s \in I^{\mathfrak{M}^\bullet}$, we have $\mathfrak{M}^\bullet, s \vDash \psi$ iff $\mathfrak{M}, s \Vdash \psi$.*

Proof of claim 2.4.11.4: We prove it by induction on ψ. For the case of $\psi \in \mathcal{L}_{\text{LCPP}}^{\texttt{B-Free}}$, this has been shown in Claim 2.4.11.3. We only focus on the cases of $\sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q$; the other cases are straightforward.

For the case of $\psi := \sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q$, we only need to show $\mu_{\pi_i}^{\mathfrak{M}}([\![\psi_i]\!]_{\pi_i}^{\mathfrak{M}}) = \mu_{\pi_i}^{\mathfrak{M}^\bullet}([\![\psi_i]\!]^{\mathfrak{M}^\bullet|^{\pi_i}})$. Please note that there is no probability formula occurring in ψ_i since $d(\varphi) = 0$. We have the following.

$$\begin{aligned}
\mu_{\pi_i}^{\mathfrak{M}}([\![\psi_i]\!]_{\pi_i}^{\mathfrak{M}}) &= \sum_{\{t \in I^{\mathfrak{M}}|^{\pi_i} \mid \mathfrak{M}, t \Vdash \psi_i\}} \mu_{\pi_i}^{\mathfrak{M}}(t) \\
&= \sum_{\{t \in I^{\mathfrak{M}}|^{\pi_i} \mid \mathfrak{M}, t \Vdash \psi_i\}} \mu_{\pi_i}^{\mathfrak{M}^\bullet}([\sigma]_t^{\pi_i}) && \text{(by Claim 2.4.11.2)} \\
&= \sum_{t \in I^{\mathfrak{M}}|^{\pi_i}} \sum_{\{\sigma' \in [\sigma]_t^{\pi_i} \mid \mathfrak{M}, t \Vdash \psi_i\}} \mu_{\pi_i}^{\mathfrak{M}^\bullet}(\sigma') \\
&= \sum_{t \in I^{\mathfrak{M}}|^{\pi_i}} \sum_{\{\sigma' \in [\sigma]_t^{\pi_i} \mid \mathfrak{M}^\bullet|^{\pi_i}, \sigma' \vDash \psi_i\}} \mu_{\pi_i}^{\mathfrak{M}^\bullet}(\sigma') && \text{(by Claim 2.4.11.3)} \\
&= \sum_{\{\sigma \in I^{\mathfrak{M}^\bullet}|^{\pi_i} \mid \mathfrak{M}^\bullet|^{\pi_i}, \sigma \vDash \psi_i\}} \mu_{\pi_i}^{\mathfrak{M}^\bullet}(\sigma) \\
&= \mu_{\pi_i}^{\mathfrak{M}^\bullet}([\![\psi_i]\!]^{\mathfrak{M}^\bullet|^{\pi_i}})
\end{aligned}$$

■

□

2.4.3 Canonical nonstandard model

Up to now, we have shown that each $\varphi \in \mathcal{L}_{\text{LCPP}}$ can be reduced to a formula $\varphi' \in \mathcal{L}_{\text{LCPP}}^{\text{B}}$ (Proposition 2.4.8) and if a formula $\varphi' \in \mathcal{L}_{\text{LCPP}}^{\text{B}}$ is satisfiable in nonstandard models, it is satisfiable in standard models (Proposition 2.4.11). To show that $\mathbb{SLCPP}$ is complete with respect to standard models, we only need to show that each $\mathbb{SLCPP}$-consistent formula $\varphi \in \mathcal{L}_{\text{LCPP}}^{\text{B}}$ is satisfiable in nonstandard models.

Let $\varphi \in \mathcal{L}_{\text{LCPP}}^{\text{B}}$ be consistent. Next, we will construct a canonical nonstandard model for φ and show the truth lemma. The canonical model will be built by levels, and the number of its levels will be bounded by the modal depth of φ. The notion of modal depth is defined in the following.

Definition 2.4.12 (Modal Depth) *The modal depth of a formula $\psi \in \mathcal{L}_{LCPP}$, denoted as $md(\varphi)$, is defined as follows.*

$$\begin{array}{lcl} md(p) & = & 0 \\ md(\neg\psi) & = & md(\psi) \\ md(\psi \wedge \psi') & = & max\{md(\psi), md(\psi')\} \\ md([a]\psi) & = & 1 + md(\psi) \\ md(\sum_{i=1}^{n} q_i B_{\pi_i}\psi_i \geq q) & = & max\{|\pi_i| + md(\psi_i) \mid 1 \leq i \leq n\} \end{array}$$

where $|\pi_i|$ is the length of the sequence π_i.

Here are some notions before we construct the model for φ. We use $\mathbf{Act}|_\varphi$ to denote the set of actions occurring in φ, and $(\mathbf{Act}|_\varphi)^n$ to denote the set of sequences whose length is no bigger than n and whose actions are in $\mathbf{Act}|_\varphi$. If s is a finite set of formulas, we use φ_s to denote $\bigwedge_{\psi\in s}\psi$. Let $\sim\psi = \chi$ if $\psi = \neg\chi$, otherwise, $\sim\psi = \neg\psi$. It is obvious that $\vdash \neg\psi \leftrightarrow \sim\psi$. We use $Sub^+(\varphi)$ to denote the set $Sub(\varphi) \cup \{\sim\psi \mid \psi \in Sub(\varphi)\}$, where $Sub(\varphi)$ is the set of all subformulas of φ.

Let $md(\varphi) = h$. Next, we will define the canonical nonstandard model for φ. Note that it is no harm to assume $h > 0$ since $\vdash \varphi \leftrightarrow \varphi \wedge [a]\top$.

Definition 2.4.13 Γ_k^φ *and $Atom_k^\varphi$ (where $0 \leq k \leq h$) are defined as follows.*

- $k = h$
  - $\Gamma_h^\varphi = \{\psi \in sub^+(\varphi) \mid md(\psi) = 0, \psi \in \mathcal{L}_{LCPP}^{B\text{-}Free}\}$;
  - $Atom_h^\varphi = \{s \mid s$ *is a maximally consistent subset of* $\Gamma_0^\varphi\}$;
- $k < h$ *but* $k > 0$
  - $\Gamma_k^\varphi = \{\psi \in sub^+(\varphi) \mid md(\psi) \leq h - k, \psi \in \mathcal{L}_{LCPP}^{B\text{-}Free}\}$
    $\cup \{sub^+(\langle a\rangle\varphi_s) \mid a \in (\mathbf{Act}|_\varphi), s \in Atom_{k+1}^\varphi\}$;
  - $Atom_k^\varphi = \{s \mid s$ *is a maximally consistent subset of* $\Gamma_k^\varphi\}$;
- $k = 0$
  - $\Gamma_0^\varphi = sub^+(\varphi)$
    $\cup \{sub^+(B_\epsilon(\psi_1 \wedge \cdots \wedge \psi_j) \geq 0) \mid \psi_1, \cdots, \psi_j \in sub^+(\varphi) \cap \mathcal{L}_{LCPP}^{B\text{-}Free}\}$
    $\cup \{sub^+(B_\epsilon\langle\pi\rangle\varphi_s \leq 0) \mid s \in Atom_k^\varphi, \pi \in (\mathbf{Act}|_\varphi)^k, 1 \leq k \leq h\}$;
  - $Atom_0^\varphi = \{s \mid s$ *is a maximally consistent subset of* $\Gamma_h^\varphi\}$

By induction on k, it is easy to show that all Γ_k^φ and all $Atom_k^\varphi$ are finite. Since each $Atom_k^\varphi$ is the set of all maximally consistent subset of Γ_k^φ, we have the following two propositions.

Proposition 2.4.14 *For each $0 \leq k \leq h$, we have* $\vdash \bigvee_{s\in Atom_k^\varphi}\varphi_s$

Proposition 2.4.15 *For each $\psi \in \Gamma_k^\varphi$, we have* $\vdash \psi \leftrightarrow \bigvee_{\{s\in Atom_k^\varphi \mid \psi\in s\}}\varphi_s$

Let u be a set in $Atom_h^\varphi$ such that $\varphi \in u$. Please note it follows by Lindenbaum's lemma that such a set u exists.

Definition 2.4.16 (Canonical Nonstandard Model) *The canonical model $\mathfrak{M}^{\varphi}_u$ is defined as follows:*

- $S^{\mathfrak{M}^{\varphi}_u} = \{(s,k) \mid s \in Atom^{\varphi}_k, 0 \leq k \leq h\}$
- $R^{\mathfrak{M}^{\varphi}_u} = \{((s,k), a, (t, k+1)) \mid \varphi_s \wedge \langle a \rangle \varphi_t \textit{ is consistent}, a \in \mathbf{Act}|_{\varphi}\}$
- $I^{\mathfrak{M}^{\varphi}_u} = \{(s,0) \mid s \textit{ and } u \textit{ contain the same probability literals}\}$
- $V^{\mathfrak{M}^{\varphi}_u}(p) = \{(s,k) \mid p \in s\}$ *for each* $p \in sub^+(\varphi)$
- $\mu_{\pi}^{\mathfrak{M}^{\varphi}_u}$ *is defined latter.*

where a probability literal is a formula of the form $\sum_{i=1}^{n} q_i \psi_i \leq q$ *or* $\neg(\sum_{i=1}^{n} q_i \psi_i \leq q)$.

Proposition 2.4.17 *For each* $(s,k) \in S^{\mathfrak{M}^{\varphi}_u}$ *and* $\langle a \rangle \psi \in sub^+(\varphi)$, $\langle a \rangle \psi \in s$ *iff there exists* $(t, k+1) \in S^{\mathfrak{M}^{\varphi}_u}$ *such that* $\psi \in t$ *and* $(s,k) \xrightarrow{a} (t, k+1)$.

PROOF We leave the proof of left-to-right to the reader. Please note that it follows by the definition that $k < h$ since $\langle a \rangle \psi \in s$ and $s \in Atom^{\varphi}_k$. Assume that $\langle a \rangle \psi \in s$ and that there does not exist $(t, k+1) \in S^{\mathfrak{M}^{\varphi}_u}$ such that $\psi \in t$ and $(s,k) \xrightarrow{a} (t, k+1)$. It follows that for all $t \in Atom^{\varphi}_{k+1}$: if $\psi \in t$ then $\varphi_s \vdash [a]\neg\varphi_t$. Let $t_1, \cdots, t_n$ be all the sets in $Atom^{\varphi}_{k+1}$ such that ψ is a member of them. It follows by Proposition 2.4.15 that $\vdash \psi \leftrightarrow \varphi_{t_1} \vee \cdots \vee \varphi_{t_n}$. Moreover, since $\vdash \varphi_s \to ([a]\neg\varphi_{t_1} \wedge \cdots \wedge [a]\varphi_{t_n})$, it is easy to show that $\vdash \varphi_s \to [a]\neg\psi$. This is in contradiction with $\langle a \rangle \psi \in s$ and the assumption that s is consistent. Therefore, we have shown if $\langle a \rangle \psi \in s$ then there exists $(t, k+1) \in S^{\mathfrak{M}^{\varphi}_u}$ such that $\psi \in t$ and $(s,k) \xrightarrow{a} (t, k+1)$. □

With the proposition above, we have the following proposition immediately, which is the "truth lemma" for formulas in $\mathcal{L}^{\texttt{B-Free}}_{\text{LCPP}}$.

Proposition 2.4.18 *For each* $\psi \in sub^+(\varphi) \cap \mathcal{L}^{B\text{-}Free}_{LCPP}$ *and each* $(s,k) \in S^{\mathfrak{M}^{\varphi}_u}$, *we have* $\mathfrak{M}^{\varphi}_u, (s,k) \Vdash \psi$ *iff* $\psi \in s$.

Next, we will focus on the probability formulas. We will show that there exist functions $\mu_{\pi}^{\mathfrak{M}^{\varphi}_u}$ to make sure the probability formulas are true.

Proposition 2.4.19 *Given* $(s,k) \in S^{\mathfrak{M}^{\varphi}_u}$ *and* $\pi \in (\mathbf{Act}|_{\varphi})^k$, $(s,k) \in I^{\mathfrak{M}^{\varphi}_u}|^{\pi}$ *implies* $\vdash \varphi_u \to B_{\pi}\varphi_s > 0$.

PROOF For the case of $k = 0$ and $\pi := \epsilon$, let $\psi \in s$ be a probability literal, and let $\chi := \wedge(s \setminus \{\psi\})$. By Proposition 2.3.14 and Proposition 2.3.2, it is easy to show that $\vdash B_{\epsilon}\varphi_s \leq 0 \leftrightarrow \neg\psi \vee B_{\epsilon}\chi \leq 0$. Since $(s,0) \in I^{\mathfrak{M}^{\varphi}_u}$, it follows that $\vdash \varphi_u \to \psi$. Thus, we have $\vdash \varphi_u \wedge B_{\epsilon}\varphi_s \leq 0 \to B_{\epsilon}\chi \leq 0$. Since $\vdash \varphi_s \to \chi$, it follows by Axiom T that $\vdash \varphi_s \to B_{\epsilon}\chi > 0$. By Definition 2.4.13, it follows that $B_{\epsilon}\chi > 0 \in \Gamma^{\varphi}_0$. Thus, we have $B_{\epsilon}\chi > 0 \in s$, and consequently $B_{\epsilon}\chi > 0 \in u$. Therefore, we have $\vdash \varphi_u \wedge B_{\epsilon}\varphi_s \leq 0 \to \bot$, and consequently $\vdash \varphi_u \to B_{\epsilon}\varphi_s > 0$.

For the case of $k+1$ and πa, it follows by $(s, k+1) \in I^{\mathfrak{M}^{\varphi}_u}|^{\pi a}$ that there exists $(w, 0) \in I^{\mathfrak{M}^{\varphi}_u}$ such that $(w,0) \xrightarrow{\pi a} (s, k+1)$. According to Proposition 2.3.15, by induction on π, it is easy to show that $\vdash B_{\epsilon}\langle \pi a \rangle \psi > 0 \to B_{\pi a}\psi > 0$. Since w and u

share the same probability formulas and $B_\epsilon\langle\pi a\rangle\varphi_s > 0 \in \Gamma_0^\varphi$, by Axioms T, we only need to show that $\langle\pi a\rangle\varphi_s \in w$. Next we will show it by induction on π. It is obvious for the case of a. For the case of πa, we have that $(w,0) \xrightarrow{\pi} (w',k) \xrightarrow{a} (s,k+1)$ for some $(w',k) \in S^{\mathfrak{M}_u^\varphi}$. It follows by induction on π that $\langle\pi\rangle\varphi_{w'} \in w$. Moreover, since $\langle a\rangle\varphi_s \in w'$, we have $\vdash \varphi_w \to \langle\pi\rangle\varphi_{w'}$ and $\vdash \varphi_{w'} \to \langle a\rangle\varphi_s$. Therefore, we have $\vdash \varphi_w \to \langle\pi a\rangle\varphi_s$, and consequently $\langle\pi a\rangle\varphi_s \in w$. □

Proposition 2.4.20 *If $B_\epsilon\psi > 0 \in u$ then there exists $(s,0) \in I^{\mathfrak{M}_u^\varphi}$ such that $\psi \in s$.*

PROOF Let D be the set of all the probability literals in u. We then only need to show that $D \cup \{\psi\}$ is consistent. If it is not, we have $\vdash \varphi_D \to \neg\psi$. It follows by Axiom PRTR(ϵ) that $B_\epsilon(\neg\varphi_D \vee \neg\psi) = 1$. It follows by Proposition 2.3.14 that $\vdash \neg\varphi_D \vee B_\epsilon\neg\psi = 1$. Since $\vdash \varphi_u \to \varphi_D$, we have $\vdash \varphi_u \to B_\epsilon\neg\psi = 1$. By Axioms PRTR($\epsilon$) and Proposition 2.3.3, it follows that $\vdash \varphi_u \to B_\epsilon\psi = 0$. This is in contradiction with $B_\epsilon\psi > 0 \in u$. Therefore, $D \cup \{\psi\}$ is consistent. □

Proposition 2.4.21 *If $\langle\pi\rangle\psi \in s$ for some $(s,0) \in S^{\mathfrak{M}_u^\varphi}$, there exists $(t,|\pi|) \in S^{\mathfrak{M}_u^\varphi}$ such that $(s,0) \xrightarrow{\pi} (t,|\pi|)$ and ψ is consistent with t.*

PROOF We prove it by induction on π. It is obvious if $\pi := \epsilon$. If it is πa, it follows by induction on π that there exists $(s',|\pi|) \in S^{\mathfrak{M}_u^\varphi}$ such that $(s,0) \xrightarrow{\pi} (s',|\pi|)$ and $\langle a\rangle\psi$ is consistent with s'. Next, we only need to show that there exists $(t,|\pi a|) \in S^{\mathfrak{M}_u^\varphi}$ such that $(s',|\pi|) \xrightarrow{a} (t,|\pi a|)$ and ψ is consistent with t.

We construct an appropriate $t \in Atom^\varphi_{|\pi a|}$ by forcing choices. Enumerate the formulas in $\Gamma^\varphi_{|\pi a|}$ as $\chi_1, \cdots, \chi_m$. Define D_0 to be $\{\psi\}$. Then $\varphi_{s'} \wedge \langle a\rangle\varphi_{D_0}$ is consistent. Suppose that D_j is a formula set such that $\varphi_{s'} \wedge \langle a\rangle\varphi_{D_j}$ is consistent where $0 \leq j \leq m$. Therefore,either for $D' = D_j \cup \{\chi_{j+1}\}$ or for $D' = D_j \cup \{\neg\chi_{j+1}\}$ we have that $\varphi_{s'} \wedge \langle a\rangle\varphi_{D'}$ is consistent. Choose D_{j+1} to this consistent expansion, and let t be $D_m \cap \Gamma^\varphi_{|\pi a|}$. Thus, we have $t \in Atom^\varphi_{|\pi a|}$, $\varphi_{s'} \wedge \langle a\rangle\varphi_t$ is consistent and t is consistent with ψ. Therefore, we have $(s',|\pi|) \xrightarrow{a} (t,|\pi a|)$ and $(s,0) \xrightarrow{\pi a} (t,|\pi a|)$. □

Proposition 2.4.22 *Given $(s,k) \in S^{\mathfrak{M}_u^\varphi}$ and $\pi \in (\mathbf{Act}|_\varphi)^k$, $(s,k) \notin I^{\mathfrak{M}_u^\varphi}|^\pi$ implies $\vdash \varphi_u \to B_\pi\varphi_s = 0$.*

PROOF For the case of $k = 0$ and $\pi := \epsilon$, without loss of generality, assuming $\neg\psi \in u$ and $\psi \in s$ for some probability literal $\psi \in \Gamma_0^\varphi$. Let $\chi := \wedge(s \setminus \{\psi\})$. By Proposition 2.3.14, it follows that $\vdash B_\epsilon\varphi_s > 0 \leftrightarrow \psi \wedge B_\epsilon\chi > 0$. Therefore, we have $\vdash \varphi_u \wedge B_\epsilon\varphi_s > 0 \to \bot$, and consequently $\vdash \varphi_u \to B_\epsilon\varphi_s \leq 0$. It follows by Axiom Nonneg(ϵ) that $\vdash \varphi_u \to B_\epsilon\varphi_s = 0$.

For the case of $k+1$ and πa, by Axiom Nonneg(ϵ), we only need to show $\vdash \varphi_u \to B_{\pi a}\varphi_s \leq 0$. If $\varphi_u \wedge B_{\pi a}\varphi_s > 0$ is consistent, it follows by Axiom PRF(πa) that $\varphi_u \wedge B_\epsilon\langle\pi a\rangle\varphi_s > 0$ is consistent. Since $B_\epsilon\langle\pi a\rangle\varphi_s > 0 \in \Gamma_0^\varphi$, it follows that $B_\epsilon\langle\pi a\rangle\varphi_s > 0 \in u$. It follows by Proposition 2.4.20 that $\langle\pi a\rangle\varphi_s \in w$ for some $(w,0) \in I^{\mathfrak{M}_u^\varphi}$. By Proposition 2.4.21 that there exists $(v,k+1) \in S^{\mathfrak{M}_u^\varphi}$ such that $(w,0) \xrightarrow{\pi a} (v,k+1)$ and φ_s is consistent with v. This means that $s = v$, and then $(s, k+1) \in I^{\mathfrak{M}_u^\varphi}|\pi a$. This is in contradiction with our assumption. Therefore, $\varphi_u \wedge B_{\pi a}\varphi_s > 0$ is not consistent, and consequently $\vdash \varphi_u \to B_{\pi a}\varphi_s = 0$. □

Proposition 2.4.23 $\vdash B_{\pi_i}\psi_i = \sum_{\{s \in Atom^\varphi_{|\pi_i|} | \psi_i \in s\}} B_{\pi_i}\varphi_s$ *if* $\sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q \in sub^+(\varphi)$.

PROOF It can be easily shown by Proposition 2.3.16 □

As we said before, Proposition 2.4.18 is the "truth lemma" for formulas in $\mathcal{L}^{\texttt{B-Free}}_{\text{LCPP}}$. Before we show the "truth lemma" for formulas in $\mathcal{L}^{\text{B}}_{\text{LCPP}}$, we fist need to show the following proposition, which says that there exists probability functions to ensure probability formulas. The proof idea is that based on Propositions 2.4.19 – 2.4.23, we list all the conditions that probabilities in nonstandard models need to satisfy (cf. Definition 2.4.9) and show that these condition formulas are consistent. Then, these consistent condition formulas will generated a consistent set of inequalities. Finally, by the completeness of linear inequality logic, there exists a solution for these inequalities.

Proposition 2.4.24 *There exists functions* $\mu_\pi^{\mathfrak{M}_u^\varphi} : I^{\mathfrak{M}_u^\varphi}|^\pi \to \mathbb{Q}^+$ *where* $\pi \in \mathbf{Act}^*$ *such that* $\mathfrak{M}_u^\varphi$ *is a nonstandard model and that* $\sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q \in u$ *iff* $\sum_{i=1}^n q_i \mu_{\pi_i}^{\mathfrak{M}_u^\varphi}(D_i) \geq q$ *where* $D_i = \{(s, |\pi_i|) \in I^{\mathfrak{M}_u^\varphi}|^{\pi_i} \mid \psi_i \in s\}$ *for each* $\sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q \in sub^+(\varphi)$.

PROOF Firstly, it follows from Proposition 2.4.14 that $\vdash \top \leftrightarrow \bigvee_{s \in Atom_0^\varphi} \varphi_s$. By Axioms $\texttt{PRTR}(\epsilon)$ and Proposition 2.3.16, it follows that

$$\vdash \sum_{s \in Atom_0^\varphi} B_\epsilon \varphi_s = 1 \tag{2.14}$$

By Proposition 2.4.19, for each $(s, 0) \in I^{\mathfrak{M}_u^\varphi}$, we have

$$u \vdash B_\epsilon \varphi_s > 0 \tag{2.15}$$

By Proposition 2.4.22, for each $(s, 0) \notin I^{\mathfrak{M}_u^\varphi}$, we have

$$u \vdash B_\epsilon \varphi_s = 0 \tag{2.16}$$

Secondly, it follows by Proposition 2.4.14 and 2.4.15 that $\vdash \top \leftrightarrow \bigvee_{s \in Atom^\varphi_{|\pi a|}} \varphi_s$ and $\vdash \langle a \rangle \top \leftrightarrow \bigvee_{\{s \in Atom^\varphi_{|\pi|} | \langle a \rangle \top \in s\}} \varphi_s$. By Proposition 2.3.6 and proposition 2.3.16, it follows that

$$\vdash \sum_{s \in Atom^\varphi_{|\pi a|}} B_{\pi a}\varphi_s = \sum_{\{s \in Atom^\varphi_{|\pi|} | \langle a \rangle \top \in s\}} B_\pi \varphi_s \tag{2.17}$$

By Proposition 2.4.19, for each $(s, |\pi a|) \in I^{\mathfrak{M}_u^\varphi}|^{\pi a}$, we have

$$u \vdash B_{\pi a}\varphi_s > 0 \tag{2.18}$$

By Proposition 2.4.19, for each $(s, |\pi a|) \notin I^{\mathfrak{M}_u^\varphi}|^{\pi a}$, we have

$$u \vdash B_{\pi a}\varphi_s = 0 \tag{2.19}$$

Thirdly, for each state set $E \subseteq I^{\mathfrak{M}^{\varphi}_u}|^{\pi a}$, it follows by Proposition 2.3.16 that $\vdash B_{\pi a} \bigvee_{(t,|\pi a|)\in E} \varphi_t = \sum_{(t,|\pi a|)\in E} B_{\pi a}\varphi_t$. For each $(t, |\pi a|) \in E$, it follows by Proposition 2.4.15 that $\vdash \langle a\rangle\varphi_t \leftrightarrow \bigvee_{\{s\in Atom^{\varphi}_{|\pi|} | \langle a\rangle\varphi_t \in s\}} \varphi_s$. What is more, since $\vdash \langle a\rangle(\bigvee_{(t,|\pi a|)\in E} \varphi_t) \leftrightarrow \bigvee_{(t,|\pi a|)\in E} \langle a\rangle\varphi_t$, we have that

$$\vdash \langle a\rangle(\bigvee_{(t,|\pi a|)\in E} \varphi_t) \leftrightarrow \bigvee_{\{s\in Atom^{\varphi}_{|\pi|} | \exists (t,|\pi a|)\in E: \langle a\rangle\varphi_t\in s\}} \varphi_s.$$

By Axiom $\mathtt{Add}(\pi)$, we have

$$\vdash B_\pi\langle a\rangle(\bigvee_{(t,|\pi a|)\in E} \varphi_t) = \sum_{\{s\in Atom^{\varphi}_{|\pi|} | \exists (t,|\pi a|)\in E: \langle a\rangle\varphi_t\in s\}} B_\pi\varphi_s.$$

By Axiom $\mathtt{PRF}(\pi a)$, we have $\vdash B_{\pi a} \bigvee_{(t,|\pi a|)\in E} \varphi_t \leq B_\pi\langle a\rangle(\bigvee_{(t,|\pi a|)\in E} \varphi_t)$. Therefore, we have

$$\vdash \sum_{(t,|\pi a|)\in E} B_{\pi a}\varphi_t \leq \sum_{\{s\in Atom^{\varphi}_{|\pi|} | \exists (t,|\pi a|)\in E: \langle a\rangle\varphi_t\in s\}} B_\pi\varphi_s \tag{2.20}$$

Moreover, for each set $E \subseteq I^{\mathfrak{M}^{\varphi}_u}|^{\pi a}$, if there exists $(s, |\pi|) \in I^{\mathfrak{M}^{\varphi}_u}|^{\pi}$ such that $R^{\mathfrak{M}^{\varphi}_u}(s, |\pi|) \cap E \neq \emptyset$ and $R^{\mathfrak{M}^{\varphi}_u}(s, |\pi|) \setminus E \neq \emptyset$, namely $(t, |\pi a|) \in E$ and $(t', |\pi a|) \notin E$ for some $(t, |\pi a|), (t, |\pi a|) \in R^{\mathfrak{M}^{\varphi}_u}(s, |\pi|)$, it follows that $\vdash \varphi_t \to \varphi_E$ (let $\varphi_E := \bigvee_{(t,|\pi a|)\in E} \varphi_t$) and $\vdash \varphi_{t'} \to \neg\varphi_E$. Therefore, we have $\vdash \langle a\rangle\varphi_t \wedge \langle a\rangle\varphi_{t'} \to \langle a\rangle\varphi_E \wedge \langle a\rangle\neg\varphi_E$. Since $\vdash \varphi_s \to \langle a\rangle\varphi_t \wedge \langle a\rangle\varphi_{t'}$, it follows that $\vdash \varphi_s \to \langle a\rangle\varphi_E \wedge \langle a\rangle\neg\varphi_E$. Therefore, we have $\vdash B_\pi\varphi_s \leq B_\pi(\langle a\rangle\varphi_E \wedge \langle a\rangle\neg\varphi_E)$. It follows by Proposition 2.4.19 that $u \vdash B_\pi\langle a\rangle\varphi_E \wedge \langle a\rangle\neg\varphi_E > 0$. Thus, by Proposition 2.3.7, we have $u \vdash B_{\pi a}\varphi_E < B_\pi\langle a\rangle\varphi_E$, namely $u \vdash B_{\pi a} \bigvee_{t\in E} \varphi_t < B_\pi \bigvee_{t\in E}\langle a\rangle\varphi_t$. Therefore, we have

$$u \vdash \sum_{(t,|\pi a|)\in E} B_{\pi a}\varphi_t < \sum_{\{(s,|\pi|)\in I^{\mathfrak{M}^{\varphi}_u}|^{\pi} | \exists (t,|\pi a|)\in E: \langle a\rangle\varphi_t\in s\}} B_\pi\varphi_s \tag{2.21}$$

Furthermore, for each $\chi := \sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q \in sub^+(\varphi)$, if $\chi \in u$, it follows by Proposition 2.4.23 that

$$u \vdash \sum_{i=1}^{n}(q_i \cdot \sum_{\{s\in Atom^{\varphi}_{|\pi_i|} | \psi_i\in s\}} B_{\pi_i}\varphi_s) \geq q \tag{2.22}$$

If $\chi \notin u$, we have

$$u \vdash \sum_{i=1}^{n}(q_i \cdot \sum_{\{s\in Atom^{\varphi}_{|\pi_i|} | \psi_i\in s\}} B_{\pi_i}\varphi_s) < q \tag{2.23}$$

Finally, we can now construct a set of linear inequalities by replacing $B_\pi\varphi_s$ in formulas of (2.14) to (2.23) by variables $x_{\pi(s,|\pi|)}$, which represents $\mu_\pi^{\mathfrak{M}^{\varphi}_u}(s, |\pi|)$. Since u is consistent, it follows that this set of linear inequalities is also consistent. By completeness of the linear inequality system, this inequality set has a solution. We define $\mu_\pi^{\mathfrak{M}^{\varphi}_u}$

by assigning $\mu_\pi^{\mathfrak{M}_u^\varphi}(s,|\pi|)$ the value of $x_{\pi(s,|\pi|)}$. It follows by (2.14) to (2.21) that $\mathfrak{M}_u^\varphi$ is a nonstandard model. For each $\chi := \sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q \in sub^+(\varphi)$, it follows by (2.22) and (2.23) that $\chi \in u$ iff $\sum_{i=1}^n q_i \mu_{\pi_i}^{\mathfrak{M}_u^\varphi}(\{s \in Atom_{|\pi_i|}^\varphi \mid \psi_i \in s\}) \geq q$. By (2.16) and (2.19), we have that $\mu_{\pi_i}^{\mathfrak{M}_u^\varphi}(s,|\pi|) = 0$ for each $(s,|\pi_i|) \notin I^{\mathfrak{M}_u^\varphi}|^{\pi_i}$. Therefore, we have $\mu_{\pi_i}^{\mathfrak{M}_u^\varphi}(\{s \in Atom_{|\pi_i|}^\varphi \mid \psi_i \in s\}) = \mu_{\pi_i}^{\mathfrak{M}_u^\varphi}(D_i)$. □

Proposition 2.4.25 *For each formula $\psi \in sub^+(\varphi)$ and each $(s,0) \in I^{\mathfrak{M}_u^\varphi}$, we have $\mathfrak{M}_u^\varphi,(s,0) \Vdash \psi$ iff $\psi \in s$.*

PROOF We prove it by induction on ψ. Since $\varphi \in \mathcal{L}_{\text{LCPP}}^{\text{B}}$, it follows that $sub^+(\varphi) \subset \mathcal{L}_{\text{LCPP}}^{\text{B}}$. By Proposition 2.4.18, we only need to focus on the case of $\sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q$. Please note that if $\psi := \sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q$ and $\psi \in sub^+(\varphi)$ then $\psi_i \in \mathcal{L}_{\text{LCPP}}^{\texttt{B-Free}}$. We have the followings.

$$
\begin{aligned}
& \mathfrak{M}_u^\varphi,(s,0) \Vdash \sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q \\
\iff & \sum_{i=1}^n q_i \mu_{\pi_i}^{\mathfrak{M}_u^\varphi} [\![\psi_i]\!]_{\pi_i}^{\mathfrak{M}_u^\varphi} \geq q \\
& \text{where } [\![\psi_i]\!]_{\pi_i}^{\mathfrak{M}_u^\varphi} = \{(t,k) \in I^{\mathfrak{M}_u^\varphi}|^{\pi_i} \mid \mathfrak{M}_u^\varphi,(t,k) \vDash \psi_i\} \\
& \text{(Please note that by induction on } \pi \text{ it is easy to show} \\
& \text{that } (t,k) \in I^{\mathfrak{M}_u^\varphi}|^{\pi} \text{ implies } k = |\pi|) \\
\iff & \sum_{i=1}^n q_i \mu_{\pi_i}^{\mathfrak{M}_u^\varphi} D_i \geq q \\
& \text{where } D_i = \{(t,|\pi_i|) \in I^{\mathfrak{M}_u^\varphi}|^{\pi_i} \mid \psi_i \in t\} \\
& \quad \text{(since } \psi_i \in \mathcal{L}_{\text{LCPP}}^{\texttt{B-Free}}\text{, by Proposition 2.4.18 we have } [\![\psi_i]\!]_{\pi_i}^{\mathfrak{M}_u^\varphi} = D_i) \\
\iff & \sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q \in u \qquad \text{(by Proposition 2.4.24)} \\
\iff & \sum_{i=1}^n q_i B_{\pi_i}\psi_i \geq q \in s \qquad \text{(by } (s,0) \in I^{\mathfrak{M}_u^\varphi})
\end{aligned}
$$

□

By Proposition 2.4.18 and Proposition 2.4.25, the following proposition follows immediately.

Proposition 2.4.26 *If $\varphi \in \mathcal{L}_{LCPP}^{B}$ is consistent then it is satisfiable in nonstandard models.*

Now we are ready for the completeness of $\mathbb{SLCPP}$ in standard models.

Theorem 2.4.27 (Completeness) *For each formula $\varphi \in \mathcal{L}_{LCPP}$, $\vDash \varphi$ implies $\vdash \varphi$.*

PROOF We only need to show that if $\neg\varphi$ is consistent then $\neg\varphi$ is satisfiable in standard models. If $\neg\varphi$ is consistent, it follows by Proposition 2.4.8 that there exists a formula $\varphi' \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$ such that $\vdash \neg\varphi \leftrightarrow \varphi'$ and φ' is consistent. It follows by Proposition 2.4.26 that φ' is satisfiable in nonstandard models. By Proposition 2.4.11, it follows that φ' is satisfiable in standard models. It follows by the soundness that $\neg\varphi$ is satisfiable in standard models. □

2.5 Decidability

This section will show that the problem whether a formula $\varphi \in \mathcal{L}_{\mathrm{LCPP}}$ is satisfiable in standard models is decidable. First, we show that the problem whether a formula $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$ is satisfiable in nonstandard models is decidable. Second, we show that a formula $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$ is satisfiable in standard models if and only if it is satisfiable in nonstandard models. Since each $\varphi \in \mathcal{L}_{\mathrm{LCPP}}$ can be reduced to be a formula $\varphi' \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$, thus the decidability of φ' in nonstandard models will lead to the decidability of φ in standard models.

Given $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$, we use $|\varphi|$ to denote the length of φ and use $f(|\varphi|)$ to denote the size of the canonical nonstandard model. It is obvious that $f(|\varphi|) \in \mathbb{N}$. We use $||\varphi||$ to denote the length of the longest coefficients that appear in φ.

Proposition 2.5.1 *If $\varphi \in \mathcal{L}^{B}_{LCPP}$ is satisfiable in nonstandard models then it is also satisfiable in a nonstandard model with at most $f(|\varphi|)$ states where the value assigned to each state by $\mu^{\mathfrak{M}}_{\pi}$ is a rational number with size of $O(r||\varphi|| + r\log r)$, where $r = O(|\varphi|^{|\varphi|} + 2^{f(|\varphi|)})$.*

PROOF If $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$ is satisfiable in nonstandard models, it follows by Proposition 2.4.11 that φ is satisfiable in standard models. By the soundness of $\mathbb{SLCPP}$ with respect to standard models, we have that φ is consistent. As it is shown in the proof of completeness, φ is satisfiable in the canonical nonstandard model whose size is $f(|\varphi|)$. Next, we will show that for each $t \in I^{\mathfrak{M}^{\varphi}_{u}}|^{\pi}$, $\mu^{\mathfrak{M}^{\varphi}_{u}}_{\pi}(t)$ is a rational number whose size can be bounded by $O(r||\varphi|| + r\log r)$. Please note that we only need to care about the action sequence $\pi \in (\mathbf{Act}|_{\varphi})^{md(\varphi)}$.

In the proof of Proposition 2.4.24, we know that the value of $\mu^{\mathfrak{M}^{\varphi}_{u}}_{\pi}(t)$ is determined by the system of linear inequalities listed by (2.14)–(2.23) in the proof of Proposition 2.4.24. Next, we will show how many linear inequalities are listed by (2.14)–(2.23).

By (2.22) and (2.23), for each χ of the form $\sum_{i=1}^{m} q_i\psi_i \geq q$ and $\chi \in sub^{+}(\varphi)$, there is a corresponding linear inequality. Therefore, (2.22) and (2.23) list at most $|\varphi|$ linear inequalities into the system.

(2.14)–(2.16) are the requirements that the function $\mu^{\mathfrak{M}^{\varphi}_{u}}_{\epsilon}$ needs to meet. They list 5 linear inequalities into the system. Please note that the linear inequality $x_1 + \cdots x_k = q$ is two inequalities in the system, that is, $x_1 + \cdots x_k \geq q$ and $(-1)x_1 + \cdots (-1)x_k \leq -q$.

(2.17)–(2.21) are the requirements that the function $\mu^{\mathfrak{M}^{\varphi}_{u}}_{\pi a}$ needs to meet for each $\pi a \in (\mathbf{Act}|_{\varphi})^{md(\varphi)}$. Given $\pi a \in (\mathbf{Act}|_{\varphi})^{md(\varphi)}$, (2.17)–(2.19) list 5 linear inequalities. (2.20)–(2.21) list 2 linear inequalities for each $E \subseteq I^{\mathfrak{M}^{\varphi}_{u}}|^{\pi a}$. Since $I^{\mathfrak{M}^{\varphi}_{u}}|^{\pi a} \subseteq S^{\mathfrak{M}^{\varphi}_{u}}$ and the size of $S^{\mathfrak{M}^{\varphi}_{u}}$ is $f(|\varphi|)$, there are at most $2^{f(|\varphi|)}$ such subset E. Therefore, (2.17)–(2.21) list at most $5 + 2 \times 2^{f(|\varphi|)}$ linear inequalities for each $\pi a \in (\mathbf{Act}|_{\varphi})^{md(\varphi)}$.

Since there are at most $|\varphi|^{|\varphi|}$sequences in $(\mathbf{Act}|_\varphi)^{md(\varphi)}$, thus (2.17)–(2.21) list at most $|\varphi|^{|\varphi|}(5 + 2 \times 2^{f(|\varphi|)})$ linear inequalities in the system.

Therefore, (2.14)–(2.23) list at most $|\varphi| + 5 + |\varphi|^{|\varphi|}(5 + 2 \times 2^{f(|\varphi|)})$ linear inequalities in the system of linear inequalities. Since $|\varphi| + 5 + |\varphi|^{|\varphi|}(5 + 2 \times 2^{f(|\varphi|)}) \leq r$, there are at most r linear inequalities in the system. It follows by Theorem A.2.8 that there exists a probability function $\mu_\pi^{\mathfrak{M}}$ such that the value assigned to each state by $\mu_\pi^{\mathfrak{M}}$ is a rational number with size of $O(r||\varphi|| + r \log r)$. □

The following proposition follows immediately.

Proposition 2.5.2 *Given $\varphi \in \mathcal{L}^B_{LCPP}$, the problem whether φ is satisfiable in nonstandard models is decidable.*

Next, we show that the problem whether $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$ is satisfiable in standard models can be reduced to the problem whether φ is satisfiable in nonstandard models.

Proposition 2.5.3 *$\varphi \in \mathcal{L}^B_{LCPP}$ is satisfiable in standard models if and only if φ is satisfiable in nonstandard models.*

PROOF Due to Proposition 2.4.11, we only need to show that if $\varphi \in \mathcal{L}^{\mathrm{B}}_{\mathrm{LCPP}}$ is satisfiable in standard models then it is also satisfiable in nonstandard models. Given a standard model $\mathcal{M} = \langle S^{\mathcal{M}}, R^{\mathcal{M}}, Pr^{\mathcal{M}}, I^{\mathcal{M}}, B^{\mathcal{M}}, V^{\mathcal{M}}\rangle$ with $\mathcal{M}, s \vDash \varphi$ for some $s \in I^{\mathcal{M}}$, we define the nonstandard model $\mathcal{M}^\bullet$ as follows.

$$
\begin{aligned}
S^{\mathcal{M}^\bullet} &= S^{\mathcal{M}} \\
R^{\mathcal{M}^\bullet} &= R^{\mathcal{M}} \\
I^{\mathcal{M}^\bullet} &= I^{\mathcal{M}} \\
V^{\mathcal{M}^\bullet} &= V^{\mathcal{M}} \\
\mu_\pi^{\mathcal{M}^\bullet} &= \mu_\pi^{\mathcal{M}}
\end{aligned}
$$

Please note that $\mu_\pi^{\mathcal{M}}$ is defined in Definition 2.2.4.

First, we need to show that $\mathcal{M}^\bullet$ is indeed a nonstandard model. We need to show the following claim.

Claim 2.5.3.1 *1. $\mu_\epsilon^{\mathcal{M}^\bullet}(I^{\mathcal{M}^\bullet}) = 1$ and $\mu_\epsilon^{\mathcal{M}^\bullet}(s) > 0$ for each $s \in I^{\mathcal{M}^\bullet}$;*

2. $\mu_{\pi a}^{\mathcal{M}^\bullet}(I^{\mathcal{M}^\bullet}|^{\pi a}) = \mu_\pi^{\mathcal{M}^\bullet}(\{s \in I^{\mathcal{M}^\bullet}|^{\pi} \mid R_a^{\mathcal{M}^\bullet}(s) \neq \emptyset\})$ and $\mu_{\pi a}^{\mathcal{M}^\bullet}(t) > 0$ for each $t \in I^{\mathcal{M}^\bullet}|^{\pi a}$;

3. $\mu_{\pi a}^{\mathcal{M}^\bullet}(E) \leq \mu_\pi^{\mathcal{M}^\bullet}(\{s \in I^{\mathcal{M}^\bullet}|^{\pi} \mid \exists t \in E : s \xrightarrow{a} t\})$ for each $E \subseteq I^{\mathcal{M}^\bullet}|^{\pi a}$;

4. $\mu_{\pi a}^{\mathcal{M}^\bullet}(E) < \mu_\pi^{\mathcal{M}^\bullet}(\{s \in I^{\mathcal{M}^\bullet}|^{\pi} \mid \exists t \in E : s \xrightarrow{a} t\})$ for each $E \subseteq I^{\mathcal{M}^\bullet}|^{\pi a}$ such that $R_a^{\mathcal{M}^\bullet}(s) \cap E \neq \emptyset$ and $R_a^{\mathcal{M}^\bullet}(s) \setminus E \neq \emptyset$ for some $s \in I^{\mathcal{M}^\bullet}|^{\pi}$.

Proof of claim 2.5.3.1:

1. Since $\mu_\epsilon^{\mathcal{M}^\bullet} = B^{\mathcal{M}}$, this is obvious.

2. First, we show $\mu_{\pi a}^{\mathcal{M}^\bullet}(t) > 0$ given $t \in I^{\mathcal{M}^\bullet}|^{\pi a}$. Since $t \in I^{\mathcal{M}^\bullet}|^{\pi a} = I^{\mathcal{M}}|^{\pi a}$, it follows that there is a sequence $s_0 a_1 \cdots s_n$ such that $s_0 \in I^{\mathcal{M}}$, $s_i \xrightarrow{a_{i+1}} s_{i+1}$ for all $0 \leq i < n$, and $s_n = t$. It follows that $B^{\mathcal{M}}(s_0) > 0$, $Pr^{\mathcal{M}}(s_i, a_{i+1}, s_{i+1}) > 0$ for all $0 \leq i < n$. It follows by Definition 2.2.4 that $\mu_{\pi a}^{\mathcal{M}}(t) \geq B^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}}(s_{i-1}, a_i, s_i)$. Since $B^{\mathcal{M}}(s_0) \times \Pi_{i=1}^{n} Pr^{\mathcal{M}}(s_{i-1}, a_i, s_i) > 0$ and $\mu_{\pi a}^{\mathcal{M}^\bullet} = \mu_{\pi a}^{\mathcal{M}}$, it follows that $\mu_{\pi a}^{\mathcal{M}^\bullet}(t) > 0$.

   Second, let $D = \{s \in I^{\mathcal{M}^\bullet}|^{\pi} \mid R_a^{\mathcal{M}^\bullet}(s) \neq \emptyset\}$ then we will show $\mu_{\pi a}^{\mathcal{M}^\bullet}(I^{\mathcal{M}^\bullet}|^{\pi a}) = \mu_{\pi}^{\mathcal{M}^\bullet}(D)$. By the definition, we only need to show $\mu_{\pi a}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi a}) = \mu_{\pi}^{\mathcal{M}}(D')$ where $D' = \{s \in I^{\mathcal{M}}|^{\pi} \mid R_a^{\mathcal{M}}(s) \neq \emptyset\}$. If $I^{\mathcal{M}}|^{\pi a} = \emptyset$, it is obvious. If $I^{\mathcal{M}}|^{\pi a} \neq \emptyset$, it follows that $I^{\mathcal{M}}|^{\pi a} = [\![\top]\!]^{\mathcal{M}|^{\pi a}}$ and $D' = [\![\langle a \rangle \top]\!]^{\mathcal{M}|^{\pi}}$. By Proposition 2.2.11, it follows that $\mu_{\pi a}^{\mathcal{M}}(I^{\mathcal{M}}|^{\pi a}) = \mu_{\pi}^{\mathcal{M}}(D')$.

3. We only need to show that $\mu_{\pi a}^{\mathcal{M}}(E) \leq \mu_{\pi}^{\mathcal{M}}(\{s \in I^{\mathcal{M}}|^{\pi} \mid \exists t \in E : s \xrightarrow{a} t\})$ for each $E \subseteq I^{\mathcal{M}}|^{\pi a}$. Given $E \subseteq I^{\mathcal{M}}|^{\pi a}$, let $D = \{s \in I^{\mathcal{M}}|^{\pi} \mid \exists t \in E : s \xrightarrow{a} t\}$. If $E = \emptyset$, it is obvious. If $E \neq \emptyset$, for each $t \in E$, there exists $s \in D$ such that $s \xrightarrow{a} t$. Moreover, it follows by Definition 2.2.4 that for each $t \in E$,

$$\mu_{\pi a}^{\mathcal{M}}(t) = \sum_{\{s \in D | s \xrightarrow{a} t\}} \mu_{\pi}^{\mathcal{M}}(s) \times Pr^{\mathcal{M}}(s, a, t)$$

   We then have the following:

$$\begin{array}{ll} & \mu_{\pi a}^{\mathcal{M}}(E) \\ = & \sum_{t \in E} \mu_{\pi a}^{\mathcal{M}}(t) \\ = & \sum_{t \in E} \big(\sum_{\{s \in D | s \xrightarrow{a} t\}} \mu_{\pi}^{\mathcal{M}}(s) \times Pr^{\mathcal{M}}(s, a, t)\big) \\ = & \sum_{s \in D} \mu_{\pi}^{\mathcal{M}}(s) \times \big(\sum_{t \in (E \cap R_a^{\mathcal{M}}(s))} Pr^{\mathcal{M}}(s, a, t)\big) \\ \leq & \sum_{s \in D} \mu_{\pi}^{\mathcal{M}}(s) \qquad \text{since } 0 < \sum_{t \in (E \cap R_a^{\mathcal{M}}(s))} Pr^{\mathcal{M}}(s, a, t) \leq 1 \\ = & \mu_{\pi}^{\mathcal{M}}(D) \end{array}$$

4. Given $u \in I^{\mathcal{M}}|^{\pi}$ and $E \subseteq I^{\mathcal{M}}|^{\pi a}$, there are $v, v' \in R_a^{\mathcal{M}}(u)$ such that $v \in E$ and $v' \notin E$. We need to show $\mu_{\pi a}^{\mathcal{M}}(E) < \mu_{\pi}^{\mathcal{M}}(D)$ where $D = \{s \in I^{\mathcal{M}}|^{\pi} \mid \exists t \in E : s \xrightarrow{a} t\}$. In 3. above, we have shown that $\mu_{\pi a}^{\mathcal{M}}(E) \leq \mu_{\pi}^{\mathcal{M}}(D)$ since for each $s \in D$:

$$\mu_{\pi}^{\mathcal{M}}(s) \times \big(\sum_{t \in (E \cap R_a^{\mathcal{M}}(s))} Pr^{\mathcal{M}}(s, a, t)\big) \leq \mu_{\pi}^{\mathcal{M}}(s)$$

   which is due to

$$0 < \sum_{t \in (E \cap R_a^{\mathcal{M}}(s))} Pr^{\mathcal{M}}(s, a, t) \leq 1.$$

   However, since there are $v, v' \in R_a^{\mathcal{M}}(u)$ such that $v \in E$ and $v' \notin E$, thus we have

$$0 < \sum_{t \in (E \cap R_a^{\mathcal{M}}(u))} Pr^{\mathcal{M}}(u, a, t) < 1.$$

Therefore, we have

$$\mu_\pi^{\mathcal{M}}(u) \times \big(\sum_{t \in (E \cap R_a^{\mathcal{M}}(u))} Pr^{\mathcal{M}}(u,a,t) \big) < \mu_\pi^{\mathcal{M}}(u).$$

Since $u \in D$, it follows that $\mu_{\pi a}^{\mathcal{M}}(E) < \mu_\pi^{\mathcal{M}}(D)$.

■

Second, by induction on the formula ψ, it is easy to show that $\mathcal{M}^\bullet, t \Vdash \psi$ if and only if $\mathcal{M}|^\pi, t \vDash \psi$ for each $\psi \in \mathcal{L}_{\text{LCPP}}^{\texttt{B-Free}}$, each $t \in S^{\mathcal{M}}$, and each $\pi \in \mathbf{Act}^*$ with $t \in I^{\mathcal{M}}|^\pi$.

Third, we will show $\mathcal{M}^\bullet, u \Vdash \psi$ if and only if $\mathcal{M}, u \vDash \psi$ for each $\psi \in \mathcal{L}_{\text{LCPP}}^{\texttt{B}}$ and each $u \in I^{\mathcal{M}}$. We prove it by induction on ψ. Please note that $\psi \in \mathcal{L}_{\text{LCPP}}^{\texttt{B}}$. Due to the second step, here we only focus on the case of $\sum_{i=1}^n q_i B_{\pi_i} \psi_i \geq q$. Since $\mu_{\pi_i}^{\mathcal{M}} = \mu_{\pi_i}^{\mathcal{M}^\bullet}$, we only need to show $[\![\psi_i]\!]^{\mathcal{M}|^{\pi_i}} = [\![\psi_i]\!]_{\pi_i}^{\mathcal{M}^\bullet}$. Since $\psi_i \in \mathcal{L}_{\text{LCPP}}^{\texttt{B-Free}}$, it follows by the second step that $[\![\psi_i]\!]^{\mathcal{M}|^{\pi_i}} = [\![\psi_i]\!]_{\pi_i}^{\mathcal{M}^\bullet}$. □

Now, we are ready to show the decidability of LCPP in standard models.

Theorem 2.5.4 (Decidability) *Given $\varphi \in \mathcal{L}_{LCPP}$, the problem whether φ is satisfiable in standard models is decidable.*

PROOF Assume the nesting degree of φ is $d(\varphi) = k$. It is obvious that $k \leq |\varphi|$. Let $\varphi_1, \cdots, \varphi_i$ where $i \leq k$ be the subformulas of φ such that $d(\varphi_j) = 1$ for all $1 \leq j \leq i$. The proofs of Proposition 2.4.6 and Proposition 2.4.7 supply procedures to reduce each φ_j to a formula φ'_j such that $\vdash \varphi_j \leftrightarrow \varphi'_j$ and $d(\varphi'_j) = 0$. Since the length of each φ_j is finite, the procedures can be terminated in a finite number of steps. We can then obtain the formula φ' by replacing each ψ_j with ψ'_j. It follows that $d(\varphi') = k - 1$. By Proposition 2.3.4, we have $\vdash \varphi \leftrightarrow \varphi'$. If $k - 1 > 0$, we do the same procedure for φ'. Therefore, we can obtain a formula ψ in a finite number of steps such that $\vdash \varphi \leftrightarrow \psi$ and $d(\psi) = 0$.

It follows by the soundness that φ is satisfiable in standard models if and only if ψ is satisfiable in standard models. Since $\psi \in \mathcal{L}_{\text{LCPP}}^{\texttt{B}}$, it follows by Proposition 2.5.3 that ψ is satisfiable in standard models if and only if it is satisfiable in nonstandard models. By Proposition 2.5.2, the problem whether ψ is satisfiable in nonstandard models is decidable. Therefore, the problem whether φ is satisfiable in standard models is decidable. □

2.6 Conclusion

In this chapter, we developed a logical framework for conformant probabilistic planning. As we argue, this approach differs from existing approaches to conformant probabilistic planning by focusing on a logical language with which to specify plans. Rather than thinking of goals of plans as subsets of the set of nodes of a probabilistic transition system, our framework allows one to think of the goal as a formula, which may be more convenient when we formulate goals that are probabilistic in nature. (Say: some message should arrive with probability greater than 0.9.) We believe that this contribution to the field of planning is very much worthwhile.

The particular logic we developed allows for reasoning about conformant plans and their probabilistic consequences. We provided an intuitive semantics, which makes it clear how probabilities change as actions are undertaken. We also provided a complete axiomatization of the logic, which shows that it is rather well-behaved for a logic that deals with conformant probabilistic planning.

In the future, we hope to expand this work so it can deal with multi-agent scenarios, where different agents may have different pieces of information about the state of the transition system and where different agents may not share the same prior probability distribution over the set of nodes of the transition system. We also hope to implement (part of) the framework so that one can actually use it for planning, even though initially we will only deal with toy examples. Still, even though these will be toy examples, we believe the addition of a logical language to a conformant probabilistic planning scenario, will already show its use in those cases.

Another direction for future research is to generalize the models by allowing states to be partially observable through sensors that map the true state of the world onto observable tokens, which is called Partially Observable Markov Decision Processes (POMDPs). For POMDP planning, the plan is usually a policy mapping of belief states onto actions (cf. (Geffner and Bonet, 2013)). Therefore, to deal with the reasoning in POMDP planning, we also need to expand the language in order to be capable of talking about policies.

Furthermore, with a possible implementation in mind, future research will include determining the complexity of algorithms for model checking and planning, which will make a comparison with standard AI approaches to planning feasible.

Chapter 3

Knowing how with intermediate constraints[1]

3.1 Introduction

Standard epistemic logic proposed by Hintikka (1962) studies propositional knowledge expressed by “knowing that φ”. However, there are very natural knowledge expressions beyond “knowing that”, such as “knowing what your password is”, “knowing why he came late”, “knowing how to go to Beijing”, and so on. In particular, *knowing how* received much attention in both philosophy and AI.

In philosophy, researchers debate about whether knowledge-how is also propositional knowledge (cf. Fantl (2008)). In AI, dating back to McCarthy and Hayes (1969), McCarthy (1979), and Moore (1985), people already started to look at it in the setting of logics of knowledge and action. However, there still is 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) and simply combining the existing modalities for “knowing that” and “ability” in a logical language like ATEL (see van der Hoek and Wooldridge (2003)) does not lead to a genuine notion of “knowing how”.

In Wang (2015a, 2016), a new approach is proposed by introducing a single new modality $\mathcal{K}h$ of goal-directed knowing how, which includes formulas $\mathcal{K}h(\psi, \varphi)$ to express that the agent knows how to achieve states in which φ is true given a precondition that ψ. The models are labelled transition systems which represent the agent’s abilities, inspired by Wang (2015b). Borrowing the idea from conformant planning in AI (cf. Smith and Weld (1998); Yu et al. (2016)), $\mathcal{K}h(\psi, \varphi)$ holds globally in a labelled transition system if there is a plan such that from all the ψ-states this plan can always be successfully executed to reach some φ-states. As an example, Figure 3.1 depicts a model, where s_i are places connected by corridors (r) or stairs (u).[2] The formula $\mathcal{K}h(p, q)$ holds in this model, since there is a plan ru which can always work to reach a q-state from any p-state. In Wang (2015a), a sound and complete proof system is given,

[1]This is an extended version of joint work with Yanjing Wang (Li and Wang (2017)).
[2]This is a variant of the running example used in Wang and Li (2012).

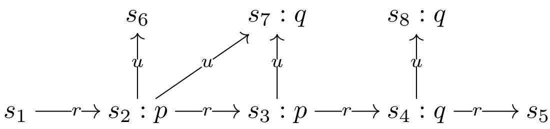


Figure 3.1

featuring a crucial axiom capturing the compositionality of plans:

$$\texttt{COMPKh} \qquad (\mathcal{K}h(p,r) \wedge \mathcal{K}h(r,q)) \to \mathcal{K}h(p,q)$$

However, as observed in Lau and Wang (2016), constraints on how we achieve the goal often matter. For example, the ways for me to go to New York are constrained by the money I have; we want to know how to win the game by playing fairly; people want to know how to be rich without breaking the law. Generally speaking, actions have costs, both financially and morally. We need to stay within our "budget" in reaching our goals. Apparently, such intermediate constraints cannot be expressed by $\mathcal{K}h(\psi,\varphi)$ since it only cares about the starting and ending states. This motivates us to introduce a ternary modality $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi)$ where χ constrains the intermediate states.

In the rest of the chapter, we first introduce the language, semantics, and a proof system of our logic in Section 3.2; in Section 3.3, we give a non-trivial completeness proof of the proof system; in Section 3.4, we show that our logic is decidable; in the last section, we conclude with future directions.

3.2 The logic KHM

This section will introduce the logic of knowing how with intermediate constraints, and we denote the logic as KHM.

3.2.1 Syntax and semantics

Firstly, we introduce the language of KHM. Besides the common boolean operators, there is a ternary modality to express knowing how and the intermediate constraints. This ternary modality was first proposed and discussed briefly in Wang (2016).

Definition 3.2.1 (Language) *Given a countable set of proposition letters* $\mathbf{P}$*, the language* $\mathcal{L}_{KHM}$ *of KHM is defined as follows:*

$$\varphi := \bot \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathcal{K}h^{\mathtt{m}}(\varphi,\varphi,\varphi)$$

where $p \in \mathbf{P}$*. We will often omit parentheses around expressions when doing so ought not cause confusion. We use the standard abbreviations* $\top, \varphi \vee \psi$ *and* $\varphi \to \psi$*, and define* $\mathcal{U}\varphi$ *as* $\mathcal{K}h^{\mathtt{m}}(\neg\varphi,\top,\bot)$*.* $\mathcal{U}$ *is intended to be an universal modality, and it will become clear after defining the semantics.*

$\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi)$ expresses that the agent knows how to guarantee φ given ψ while taking a route that satisfies χ (excluding the start and the end). Note that the formula

$\mathcal{K}h^{\mathtt{m}}(\psi \wedge \chi, \chi, \varphi \wedge \chi)$ expresses knowing how with inclusive intermediate constraints. Note that the binary know-how operator in Wang (2015b) can be defined as $\mathcal{K}h(\psi, \varphi) := \mathcal{K}h^{\mathtt{m}}(\psi, \top, \varphi)$.

The language is interpreted on models which are labelled transition systems. The model illustrates what actions the agent can do in each state.

Definition 3.2.2 (Model) *Given a countable set of proposition letters* $\mathbf{P}$*, a model is essentially a labelled transition system* (S, Act, R, V) *where:*

- S *is a non-empty set of states;*
- *Act is a set of actions;*
- $R : Act \to 2^{S\times S}$ *is a collection of transitions labelled by actions in Act;*
- $V : S \to 2^{\mathbf{P}}$ *is a valuation function.*

We write $s \xrightarrow{a} t$ *if* $(s, t) \in R(a)$*. For a sequence* $\sigma = a_1 \dots a_n \in Act^*$ *(Act^* is all the finite sequence generated by actions in Act), we write* $s \xrightarrow{\sigma} t$ *if there exist* $s_2 \dots s_n$ *such that* $s \xrightarrow{a_1} s_2 \xrightarrow{a_2} \cdots \xrightarrow{a_{n-1}} s_n \xrightarrow{a_n} t$*. Note that* σ *can be the empty sequence* ϵ *(when* $n = 0$*), and we set* $s \xrightarrow{\epsilon} s$ *for all* s*. Let* σ_k *be the initial segment of* σ *up to* a_k *for* $k \leq |\sigma|$*. In particular let* $\sigma_0 = \epsilon$*.*

Note that the labels in *Act* do not appear in the language. The graph in Figure 3.1 represents a model. We also call an action sequence a *plan*. Normally, we say that a plan σ is *executable* in a state s if there exists a state $t \in S$ such that $s \xrightarrow{\sigma} t$, which means the agent can do σ in the state s. Next, we define a notion "strongly executable" which means the agent will never fail if she performs a plan in a state.

Definition 3.2.3 (Strongly executable) *We say* $\sigma = a_1 \cdots a_n$ *is strongly executable at* s' *if for each* $0 \leq k < n$*:* $s' \xrightarrow{\sigma_k} t$ *implies that t has at least one* a_{k+1}*-successor.*

Intuitively, σ is *strongly executable* at s if you can always successfully finish the whole σ after executing any initial segment of σ from s. For example, ab is not strongly executable at s_1 in the model below, though it is executable at s_1.

$$\begin{tikzcd}
 & s_2 \arrow[r, "b"] & s_4 : q \\
s_1 : p \arrow[ur, "a"] \arrow[dr, "a"] & & \\
 & s_3 &
\end{tikzcd}$$

Definition 3.2.4 (Semantics) *Suppose* s *is a state in a model* $\mathcal{M} = (S, Act, R, V)$*, we then inductively define the notion of a formula* φ *being satisfied (or true) in* $\mathcal{M}$ *at state* s *as follows:*

$\mathcal{M}, s \vDash \bot$		*never*
$\mathcal{M}, s \vDash p$	$\iff$	$p \in V(s)$.
$\mathcal{M}, s \vDash \neg\varphi$	$\iff$	$\mathcal{M}, s \nvDash \varphi$.
$\mathcal{M}, s \vDash \varphi \wedge \psi$	$\iff$	$\mathcal{M}, s \vDash \varphi$ *and* $\mathcal{M}, s \vDash \psi$.
$\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$	$\iff$	*there exists a* $\sigma \in Act^*$ *such that for each* s' *with* $\mathcal{M}, s' \vDash \psi$*,* σ *is strongly* χ*-executable at* s' *and* $\mathcal{M}, t \vDash \varphi$ *for all* t *with* $s' \xrightarrow{\sigma} t$*,*

where we say $\sigma = a_1 \cdots a_n$ is strongly χ-executable at s' if:

- *σ is strongly executable at s', and*
- *$s' \xrightarrow{\sigma_k} t$ implies $\mathcal{M}, t \vDash \chi$ for all $0 < k < n$.*

It is obvious that ϵ is strongly χ-executable at each state s for each formula χ. Note that $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ expresses that there is a $\sigma \in Act \cup \{\epsilon\}$ such that the agent knows doing σ in ψ-states can guarantee φ, namely the witness plan σ is at most one-step. As an example, $\mathcal{K}h^{\mathtt{m}}(p, \bot, o)$ and $\mathcal{K}h^{\mathtt{m}}(p, o, q)$ hold in the following model for the witness plans a and ab respectively. Note that the truth value of $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ does not depend on the designated state.

$$\begin{tikzcd}
 & s_2 : o \arrow[dr, "b"] & \\
s_1 : p \arrow[ur, "a"] \arrow[dr, "b"'] & & s_4 : q \\
 & s_3 : \neg o \arrow[ur, "a"'] &
\end{tikzcd}$$

Now we can also check that the operator $\mathcal{U}$ defined by $\mathcal{K}h^{\mathtt{m}}(\neg\psi, \top, \bot)$ is indeed a *universal modality*:

$$\boxed{\mathcal{M}, s \vDash \mathcal{U}\varphi \quad \Leftrightarrow \quad \text{for all } t \in S, \mathcal{M}, t \vDash \varphi}$$

3.2.2 A deductive system

In this subsection, we provide a Hilbert-style proof system for the logic KHM. A proof consists of a sequence of formulas such that each formula either is an instance of an axiom or can be obtained by applying one of the rules to formulas occurring earlier in the sequence.

Definition 3.2.5 (Deductive System $\mathbb{SKHM}$) *The axioms and rules shown in Table 3.1 constitute the proof system $\mathbb{SKHM}$. We write $\mathbb{SKHM} \vdash \varphi$ (or sometimes just $\vdash \varphi$) to mean that the formula φ is derivable in the axiomatic system $\mathbb{SKHM}$; the negation of $\mathbb{SKHM} \vdash \varphi$ is written $\mathbb{SKHM} \nvdash \varphi$ (or just $\nvdash \varphi$). To say that a set D of formulas is $\mathbb{SKHM}$-inconsistent (or just inconsistent) means that there is a finite subset $D' \subseteq D$ such that $\vdash \neg \bigwedge D'$, where $\bigwedge D' := \bigwedge_{\varphi \in D'} \varphi$ if $D' \neq \emptyset$ and $\bigwedge_{\varphi \in \emptyset} \varphi := \top$. To say that a set of formulas is $\mathbb{SKHM}$-consistent (or just consistent) means that the set of formulas is not inconsistent. Consistency or inconsistency of a formula refers to the consistency or inconsistency of the singleton set containing the formula.*

Note that `DISTU`, `NECU`, `TU` are standard for the universal modality $\mathcal{U}$. `4KhmU` and `4KhmU` are introspection axioms reflecting that $\mathcal{K}h^{\mathtt{m}}$ formulas are global. `EMPKhm` captures the interaction between $\mathcal{U}$ and $\mathcal{K}h^{\mathtt{m}}$ via the empty plan. `COMPKhm` is the new composition axiom for $\mathcal{K}h^{\mathtt{m}}$. `UKhm` shows how we can weaken the knowing how claims. `ONEKhm` is the characteristic axiom for $\mathbb{SKHM}$ compared to the system for binary $\mathcal{K}h$, and it expresses the condition for the necessity of the intermediate steps.

Remark 3.2.6 *Note that the corresponding axioms for* `COMPKhm`, `EMPKhm` *and* `UKhm` *in the setting of binary $\mathcal{K}h$ are the following:*

`COMPKh`	$\mathcal{K}h(p, q) \wedge \mathcal{K}h(q, r) \to \mathcal{K}h(p, r)$
`EMPKh`	$\mathcal{U}(p \to q) \to \mathcal{K}h(p, q)$
`UKh`	$\mathcal{U}(p' \to p) \wedge \mathcal{U}(q \to q') \wedge \mathcal{K}h(p, q) \to \mathcal{K}h(p', q')$

Axioms	
`TAUT`	all tautologies of propositional logic
`DISTU`	$(\mathcal{U}p \land \mathcal{U}(p \to q)) \to \mathcal{U}q$
`TU`	$\mathcal{U}p \to p$
`4KhmU`	$\mathcal{K}h^{\mathtt{m}}(p,o,q) \to \mathcal{U}\mathcal{K}h^{\mathtt{m}}(p,o,q)$
`5KhmU`	$\neg\mathcal{K}h^{\mathtt{m}}(p,o,q) \to \mathcal{U}\neg\mathcal{K}h^{\mathtt{m}}(p,o,q)$
`EMPKhm`	$\mathcal{U}(p \to q) \to \mathcal{K}h^{\mathtt{m}}(p,\bot,q)$
`COMPKhm`	$(\mathcal{K}h^{\mathtt{m}}(p,o,r) \land \mathcal{K}h^{\mathtt{m}}(r,o,q) \land \mathcal{U}(r \to o)) \to \mathcal{K}h^{\mathtt{m}}(p,o,q)$
`ONEKhm`	$(\mathcal{K}h^{\mathtt{m}}(p,o,q) \land \neg\mathcal{K}h^{\mathtt{m}}(p,\bot,q)) \to \mathcal{K}h^{\mathtt{m}}(p,\bot,o)$
`UKhm`	$(\mathcal{U}(p' \to p) \land \mathcal{U}(o \to o') \land \mathcal{U}(q \to q') \land \mathcal{K}h^{\mathtt{m}}(p,o,q)) \to \mathcal{K}h^{\mathtt{m}}(p',o',q')$
Rules	
`MP` $\dfrac{\varphi, \varphi \to \psi}{\psi}$	`NECU` $\dfrac{\varphi}{\mathcal{U}\varphi}$ `SUB` $\dfrac{\varphi(p)}{\varphi[\psi/p]}$

Table 3.1: System $\mathbb{SKHM}$

In the system $\mathbb{SKH}$ of Wang (2015a), `UKh` *(which is called* `WKKh` *there) can be derived using* `COMPKh` *and* `EMPKh`*. However,* `UKhm` *cannot be derived using* `COMPKhm` *and* `EMPKhm`*. In particular, $\mathcal{K}h^{\mathtt{m}}(p',\bot,p) \land \mathcal{K}h^{\mathtt{m}}(p,o,q) \to \mathcal{K}h^{\mathtt{m}}(p',o,q)$ is not valid due to the lack of $\mathcal{U}(p \to o)$, in contrast with the $\mathbb{SKH}$-derivable $\mathcal{K}h(p',p) \land \mathcal{K}h(p,q) \to \mathcal{K}h(p',q)$ which is crucial in the derivation of* `UKh` *in $\mathbb{SKH}$.*

Below we derive some theorems and rules that are useful in the later proofs.

Proposition 3.2.7 *We can derive the following in $\mathbb{SKHM}$:*

`4U`	$\mathcal{U}p \to \mathcal{U}\mathcal{U}p$
`5U`	$\neg\mathcal{U}p \to \mathcal{U}\neg\mathcal{U}p$
`ULKhm`	$(\mathcal{U}(p' \to p) \land \mathcal{K}h^{\mathtt{m}}(p,o,q)) \to \mathcal{K}h^{\mathtt{m}}(p',o,q)$
`UMKhm`	$(\mathcal{U}(o \to o') \land \mathcal{K}h^{\mathtt{m}}(p,o,q)) \to \mathcal{K}h^{\mathtt{m}}(p,o',q)$
`URKhm`	$(\mathcal{U}(q \to q') \land \mathcal{K}h^{\mathtt{m}}(p,o,q')) \to \mathcal{K}h^{\mathtt{m}}(p,o,q')$
`UNIV`	$\mathcal{U}\neg p \to \mathcal{K}h^{\mathtt{m}}(p,\bot,\bot)$
`REU`	*from* $\varphi \leftrightarrow \psi$ *prove* $\mathcal{U}\varphi \leftrightarrow \mathcal{U}\psi$
`RE`	*from* $\varphi \leftrightarrow \psi$ *prove* $\chi \leftrightarrow \chi'$ *where* χ' *is obtained by replacing some occurrences of* φ *in* χ *by* ψ*.*

PROOF `REU` is immediate given `DISTU` and `NECU`. `4U` and `5U` are special cases of `4KhmU` and `5KhmU` respectively. `ULKhm`, `UMKhm`, `URKhm` are the special cases of `UKhm`. To prove `UNIV`, first, note that $\mathcal{U}\neg p \leftrightarrow \mathcal{U}(p \to \bot)$ due to `REU`. Then due to `EMPKhm`, we have $\mathcal{U}\neg p \to \mathcal{K}h^{\mathtt{m}}(p,\bot,\bot)$. `RE` can be obtained by an inductive proof on the shape of χ, which uses `UKhm` and `NECU` for the case of $\mathcal{K}h^{\mathtt{m}}(\cdot,\cdot,\cdot)$. □

Next, we will show that $\mathbb{SKHM}$ is sound with respect to the semantics provided in Section 3.2.1. Firstly we show that axioms `EMPKhm`, `COMPKhm`, `ONEKhm`, and `UKhm` are valid.

Proposition 3.2.8 $\vDash \mathcal{U}(p \to q) \to \mathcal{K}h^{\mathtt{m}}(p, \bot, q)$

PROOF Assuming that $\mathcal{M}, s \vDash \mathcal{U}(p \to q)$, it means that $\mathcal{M}, t \vDash p \to q$ for all $t \in S$. Given $\mathcal{M}, t \vDash p$, it follows that $\mathcal{M}, t \vDash q$. Thus, we have ϵ is strongly $\bot$-executable at t. Therefore, we have $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, \bot, q)$. □

Proposition 3.2.9 $\vDash \mathcal{K}h^{\mathtt{m}}(p, o, r) \wedge \mathcal{K}h^{\mathtt{m}}(r, o, q) \wedge \mathcal{U}(r \to o) \to \mathcal{K}h^{\mathtt{m}}(p, o, q)$

PROOF Assuming $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, o, r) \wedge \mathcal{K}h^{\mathtt{m}}(r, o, q) \wedge \mathcal{U}(r \to o)$, we will show that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, o, q)$. Since $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, o, r)$, it follows that there exists $\sigma \in \mathit{Act}^*$ such that for each $\mathcal{M}, u \vDash p$, σ is strongly o-executable at u and that $\mathcal{M}, v \vDash r$ for each v with $u \xrightarrow{\sigma} v$. Since $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(r, o, q)$, it follows that there exists $\sigma' \in \mathit{Act}^*$ such that for each $\mathcal{M}, v' \vDash r$, σ' is strongly o-executable at v' and that $\mathcal{M}, t \vDash q$ for each t with $v' \xrightarrow{\sigma} t$. In order to show $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, o, q)$, we only need to show that $\sigma\sigma'$ is strongly o-executable at u and that $\mathcal{M}, t' \vDash q$ for each t' with $u \xrightarrow{\sigma\sigma'} t'$, where u is a state with $\mathcal{M}, u \vDash p$.

By assumption, we know that σ is strongly o-executable at u, and for each v with $u \xrightarrow{\sigma} v$, it follows by assumption that $\mathcal{M}, v \vDash r$ and σ' is strongly o-executable at v. Moreover, since $\mathcal{M}, s \vDash \mathcal{U}(r \to o)$, it follows that $\mathcal{M}, v \vDash o$ for each v with $u \xrightarrow{\sigma} v$. Thus, $\sigma\sigma'$ is strongly o-executable at u. What is more, for each t' with $u \xrightarrow{\sigma\sigma'} t'$, there is v such that $u \xrightarrow{\sigma} v \xrightarrow{\sigma'} t'$ and $\mathcal{M}, v \vDash r$, it follows by assumption that $\mathcal{M}, t' \vDash q$. Therefore, we have $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, o, q)$. □

Proposition 3.2.10 $\vDash \mathcal{K}h^{\mathtt{m}}(p, o, q) \wedge \neg\mathcal{K}h^{\mathtt{m}}(p, \bot, q) \to \mathcal{K}h^{\mathtt{m}}(p, \bot, o)$

PROOF Assuming $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, o, q) \wedge \neg\mathcal{K}h^{\mathtt{m}}(p, \bot, q)$, we will show that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, \bot, o)$. Since $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, o, q)$, it follows that there exists $\sigma \in \mathit{Act}^*$ such that for each $\mathcal{M}, u \vDash p$, σ is strongly o-executable at u and $\mathcal{M}, v \vDash q$ for all v with $u \xrightarrow{\sigma} v$. If $\sigma \in \mathit{Act} \cup \{\epsilon\}$, it follows that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, \bot, q)$. Since $\mathcal{M}, s \vDash \neg\mathcal{K}h^{\mathtt{m}}(p, \bot, q)$, it follows that $\sigma \notin \mathit{Act} \cup \{\epsilon\}$. Thus, $\sigma = a_1 \cdots a_n$ where $n \geq 2$. Let u be a state such that $\mathcal{M}, u \vDash p$. Since $\sigma = a_1 \cdots a_n$ is strongly o-executable at u, it follows that a_1 is executable at u. Moreover, since $n \geq 2$, we have $\mathcal{M}, v \vDash o$ for each v with $u \xrightarrow{a_1} v$. Therefore, we have $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, \bot, o)$. □

Proposition 3.2.11 $\vDash \mathcal{U}(p' \to p) \wedge \mathcal{U}(o \to o') \wedge \mathcal{U}(q \to q') \wedge \mathcal{K}h^{\mathtt{m}}(p, o, q) \to \mathcal{K}h^{\mathtt{m}}(p', o', q')$

PROOF Assuming $\mathcal{M}, s \vDash \mathcal{U}(p' \to p) \wedge \mathcal{U}(o \to o') \wedge \mathcal{U}(q \to q') \wedge \mathcal{K}h^{\mathtt{m}}(p, o, q)$, we will show that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p', o', q')$. Since $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{m}}(p, o, q)$, it follows that there exists $\sigma \in \mathit{Act}^*$ such that for each $\mathcal{M}, u \vDash p$: σ is strongly o-executable at u and $\mathcal{M}, v \vDash q$ for each v with $u \xrightarrow{\sigma} v$. Let s' be a state with $\mathcal{M}, s' \vDash p'$. Next we will show that σ is strongly o'-executable at s' and $\mathcal{M}, v' \vDash q'$ for all v' with $s' \xrightarrow{\sigma} v'$.

Since $\mathcal{M}, s \vDash \mathcal{U}(p' \to p)$, it follows that $\mathcal{M}, s' \vDash p$. Thus, σ is strongly o-executable at s' and $\mathcal{M}, v' \vDash q$ for each v' with $s' \xrightarrow{\sigma} v'$. Since $\mathcal{M}, s \vDash \mathcal{U}(o \to o')$, it follows that σ is strongly o'-executable at s'. Since $\mathcal{M}, s \vDash \mathcal{U}(q \to q')$, it follows that $\mathcal{M}, v' \vDash q'$ for each v' with $s' \xrightarrow{\sigma} v'$. □

Since $\mathcal{U}$ is a universal modality, `DISTU` and `TU` are obviously valid. Since the modality $\mathcal{K}h^{\mathtt{m}}$ is not local, it is easy to show that `4KhmU` and `5KhmU` are valid. Moreover, by Propositions 3.2.8–3.2.11, we have that all axioms are valid. Due to a standard argument in modal logic, we know that the rules `MP`, `NECU` and `SUB` preserve a formula's validity. The soundness of $\mathbb{SKHM}$ follows immediately.

Theorem 3.2.12 (Soundness) $\mathbb{SKHM}$ *is sound w.r.t. the class of all models.*

3.3 Deductive completeness

This section will prove that $\mathbb{SKHM}$ is complete w.r.t. the class of all models. For the same reason as in Wang (2015a), we will build a canonical model for a given maximally consistent set (MCS). The reason is that the semantics of $\mathcal{K}h^{\mathtt{m}}$ formulas does not depend on the current state. Thus if they are true, they are true everywhere in the model. It follows that we cannot build a *single* canonical model to realize all the consistent sets of $\mathcal{L}_{\text{KHM}}$ formulas simultaneously. Instead, for each maximally consistent set of $\mathcal{L}_{\text{KHM}}$ formulas we build a separate canonical model.

However, the canonical model here is much more complicated. The reason is that in the canonical model in Wang (2015a), each knowing-how formula $\mathcal{K}h(\psi, \varphi)$ is realized by a one-step witness plan, which does not work here. Some knowing-how formulas here, such as $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \wedge \neg\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$, require their witness plans to have more than one step. Therefore, we need a new method to construct the canonical model. There are two new features for the canonical model here. Firstly, the state of the canonical model is a pair consisting of a maximally consistent set and a marker which will play an important role in defining the witness plan for $\mathcal{K}h^{\mathtt{m}}$-formulas. Secondly, formulas of the form $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ are realized by one-step plans, and formulas like $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \wedge \neg\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ are realized by two-step plans. Moreover, we deal with the second step of the plan differently from the first step of the plan.These features will become clear in the following context.

Definition 3.3.1 *We say that a set Δ of formulas is* maximally consistent *in $\mathcal{L}_{KHM}$ if Δ is consistent, and any set of formulas properly containing Γ is inconsistent. If Δ is a maximally consistent set of formulas then we say it is an MCS.*

Let Γ be an MCS in $\mathcal{L}_{\text{KHM}}$. In the following, we will build a canonical model for Γ. We first prepare ourselves with some auxiliary notions and some handy propositions.

Given a set of $\mathcal{L}_{\text{KHM}}$ formulas Δ, let $\Delta|_{\mathcal{K}h^{\mathtt{m}}}$ and $\Delta|_{\neg\mathcal{K}h^{\mathtt{m}}}$ be the collections of its positive and negative $\mathcal{K}h^{\mathtt{m}}$ formulas:

$$\Delta|_{\mathcal{K}h^{\mathtt{m}}} = \{\theta \mid \theta = \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Delta\};$$

$$\Delta|_{\neg\mathcal{K}h^{\mathtt{m}}} = \{\theta \mid \theta = \neg\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Delta\}.$$

Definition 3.3.2 *Let Φ_Γ be the set of all MCS Δ such that $\Delta|_{\mathcal{K}h^{\mathtt{m}}} = \Gamma|_{\mathcal{K}h^{\mathtt{m}}}$.*

Note that Φ_Γ is the set of all MCSs that share the same $\mathcal{K}h^{\mathtt{m}}$ formulas with Γ. The canonical model for Γ will be based on the MCSs in Φ_Γ. Since every $\Delta \in \Phi_\Gamma$ is maximally consistent, the following proposition shows an obvious property of Φ_Γ.

Proposition 3.3.3 *For each $\Delta \in \Phi_\Gamma$, we have that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Gamma$ if and only if $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Delta$ for all $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \mathcal{L}_{KHM}$.*

By a standard argument of Lindenbaum's lemma (cf. Blackburn et al. (2001)), we have the following proposition.

Proposition 3.3.4 *If Δ is consistent then there is an MCS Γ such that $\Delta \subseteq \Gamma$.*

The following proposition reveals a crucial property of Φ_Γ, which will be used repeatedly later.

Proposition 3.3.5 *If $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$ then $\mathcal{U}\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$.*

PROOF Suppose $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$, then by the definition of Φ_Γ, $\neg\varphi$ is not consistent with $\Gamma|_{\mathcal{K}h^{\mathtt{m}}} \cup \Gamma|_{\neg\mathcal{K}h^{\mathtt{m}}}$, for otherwise $\Gamma|_{\mathcal{K}h^{\mathtt{m}}} \cup \Gamma|_{\neg\mathcal{K}h^{\mathtt{m}}} \cup \{\neg\varphi\}$ can be extended into a maximally consistent set in Φ_Γ due to Proposition 3.3.4. Thus there are formulas $\mathcal{K}h^{\mathtt{m}}(\psi_1, \chi_1, \varphi_1), \ldots, \mathcal{K}h^{\mathtt{m}}(\psi_k, \chi_k, \varphi_k) \in \Gamma|_{\mathcal{K}h^{\mathtt{m}}}$ and formulas $\neg\mathcal{K}h^{\mathtt{m}}(\psi'_1, \chi'_1, \varphi'_1), \ldots, \neg\mathcal{K}h^{\mathtt{m}}(\psi'_l, \chi'_l, \varphi'_l) \in \Gamma|_{\neg\mathcal{K}h^{\mathtt{m}}}$ such that

$$\vdash (\bigwedge_{1 \le i \le k} \mathcal{K}h^{\mathtt{m}}(\psi_i, \chi_i, \varphi_i) \wedge \bigwedge_{1 \le j \le l} \neg\mathcal{K}h^{\mathtt{m}}(\psi'_j, \chi'_j, \varphi'_j)) \to \varphi.$$

By NECU,

$$\vdash \mathcal{U}((\bigwedge_{1 \le i \le k} \mathcal{K}h^{\mathtt{m}}(\psi_i, \chi_i, \varphi_i) \wedge \bigwedge_{1 \le j \le l} \neg\mathcal{K}h^{\mathtt{m}}(\psi'_j, \chi'_j, \varphi'_j)) \to \varphi).$$

By DISTU we have:

$$\vdash \mathcal{U}(\bigwedge_{1 \le i \le k} \mathcal{K}h^{\mathtt{m}}(\psi_i, \chi_i, \varphi_i) \wedge \bigwedge_{1 \le j \le l} \neg\mathcal{K}h^{\mathtt{m}}(\psi'_j, \chi'_j, \varphi'_j)) \to \mathcal{U}\varphi.$$

Since $\mathcal{K}h^{\mathtt{m}}(\psi_1, \chi_1, \varphi_1), \ldots, \mathcal{K}h^{\mathtt{m}}(\psi_k, \chi_k, \varphi_k) \in \Gamma$, we have $\mathcal{U}\mathcal{K}h^{\mathtt{m}}(\psi_1, \chi_1, \varphi_1), \ldots, \mathcal{U}\mathcal{K}h^{\mathtt{m}}(\psi_k, \chi_k, \varphi_k) \in \Gamma$ due to 4KhmU and the fact that Γ is a maximally consistent set. Similarly, we have $\mathcal{U}\neg\mathcal{K}h^{\mathtt{m}}(\psi'_1, \chi'_1, \varphi'_1), \ldots, \mathcal{U}\neg\mathcal{K}h^{\mathtt{m}}(\psi'_l, \chi'_l, \varphi'_l) \in \Gamma$ due to 5KhmU. By DISTU and NECU, it is easy to show that $\vdash \mathcal{U}(p \wedge q) \leftrightarrow \mathcal{U}p \wedge \mathcal{U}q$. Then due to a slight generalization, we have:

$$\mathcal{U}(\bigwedge_{1 \le i \le k} \mathcal{K}h^{\mathtt{m}}(\psi_i, \chi_i, \varphi_i) \wedge \bigwedge_{1 \le j \le l} \neg\mathcal{K}h^{\mathtt{m}}(\psi'_j, \chi'_j, \varphi'_j)) \in \Gamma.$$

Now it is immediate that $\mathcal{U}\varphi \in \Gamma$. Due to Proposition 3.3.3, $\mathcal{U}\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. $\Box$

Proposition 3.3.6 *Given $\mathcal{K}h^{\mathtt{m}}(\psi, \top, \varphi) \in \Gamma$ and $\Delta \in \Phi_\Gamma$, if $\psi \in \Delta$ then there exists $\Delta' \in \Phi_\Gamma$ such that $\varphi \in \Delta'$.*

PROOF Assuming $\mathcal{K}h^{\mathtt{m}}(\psi, \top, \varphi) \in \Gamma$ and $\psi \in \Delta \in \Phi_\Gamma$, if there does not exist $\Delta' \in \Phi_\Gamma$ such that $\varphi \in \Delta'$, then $\neg\varphi \in \Delta'$ for all $\Delta' \in \Phi_\Gamma$. It follows by Proposition 3.3.5 that $\mathcal{U}\neg\varphi \in \Gamma$, namely $\mathcal{K}h^{\mathtt{m}}(\varphi, \top, \bot) \in \Gamma$. Since $\mathcal{U}(\varphi \to \bot)$ and $\mathcal{K}h^{\mathtt{m}}(\psi, \top, \varphi) \in \Gamma$, it follows by `COMPKhm` that $\mathcal{K}h^{\mathtt{m}}(\psi, \top, \bot) \in \Gamma$ namely, $\mathcal{U}\neg\psi \in \Gamma$. By Proposition 3.3.3, we have that $\mathcal{U}\neg\psi \in \Delta$. It follows by `TU` that $\neg\psi \in \Delta$. This is in contradiction with $\psi \in \Delta$. Therefore, there exists $\Delta' \in \Phi_\Gamma$ such that $\varphi \in \Delta'$. □

Before building the canonical model for Γ, we firstly define the set of actions that would be part of the canonical model.

Definition 3.3.7 *The set of action symbols Act_Γ is defined as below.*

$$\begin{aligned} Act_\Gamma = &\{\langle \psi, \bot, \varphi \rangle \mid \mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi) \in \Gamma\} \cup \\ &\{\langle \chi^\psi, \varphi \rangle \mid \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi), \neg\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi) \in \Gamma\} \end{aligned}$$

The first part of Act_Γ is meant to handle the formulas $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$. These formulas can be witnessed by plans whose length are at most 1, and there are no intermediate states in such plans. The second part of Act_Γ is to handle the cases where the intermediate state is indeed necessary: $\neg\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ makes sure that you cannot have a plan to guarantee φ in less than two steps.

In the following we build a separate canonical model for the MCS Γ, for it is not possible to satisfy all $\mathcal{K}h^{\mathtt{m}}$ formulas simultaneously in a single model since those formulas are evaluated globally. Even if $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ and $\neg\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ are both consistent, they cannot be true in the same model. Because the later proofs are quite technical, it is very important to first understand the ideas behind the canonical model construction. Note that to satisfy a $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ formula, there are two cases to be considered:

(1) $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ holds and we just need a one-step witness plan, which can be handled similarly using the techniques developed in Wang (2015a);

(2) $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ does not hold, and we need to have a witness plan which at least involves an intermediate χ-stage. By `ONEKhm`, $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \chi)$ holds. It is then tempting to reduce $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ to $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \chi) \wedge \mathcal{K}h^{\mathtt{m}}(\chi, \chi, \varphi)$. However, this is not correct since we may not have a strongly χ-executable plan to reach a φ-state from *every* χ-state. Note that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ and $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \chi)$ only ensure that we can reach φ-states from χ-states that result from the witness plan for $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \chi)$. However, we cannot refer to such χ-states in the language of $\mathcal{L}_{\text{KHM}}$. This is why we include χ^ψ markers in the building blocks of the canonical model. A label χ^ψ roughly tells us where this state "comes from". [3]

Definition 3.3.8 (Canonical Model) *The canonical model for Γ is a quadruple $\mathcal{M}^c_\Gamma = \langle S^c, Act_\Gamma, R^c, V^c \rangle$ where:*

- $S^c = \{(\Delta, \chi^\psi) \mid \chi \in \Delta \in \Phi_\Gamma$*, and* $(\langle \chi^\psi, \varphi \rangle \in Act_\Gamma$ *for some* φ *or* $\langle \psi, \bot, \chi \rangle \in Act_\Gamma)\}$. *We write the pair in S as $w, v, \cdots$, and refer to the first entry of $w \in S$ as $\mathtt{L}(w)$, to the second entry as $\mathtt{R}(w)$;*
- $w \xrightarrow{\langle \psi, \bot, \varphi \rangle}_c w'$ *iff* $\psi \in \mathtt{L}(w)$ *and* $\mathtt{R}(w') = \varphi^\psi$*;*

[3] In Wang (2015a), the canonical models are much simpler: the state of the canonical model is just MCS and the canonical relations are simply labelled by $\langle \psi, \varphi \rangle$ for $\mathcal{K}h(\psi, \varphi) \in \Gamma$.

- $w \xrightarrow{\langle\chi^\psi,\varphi\rangle}_c w'$ *iff* $\mathtt{R}(w) = \chi^\psi$ *and* $\varphi \in \mathtt{L}(w')$*;*
- $p \in V^c(w)$ *iff* $p \in \mathtt{L}(w)$.

For each $w \in S$*, we also call* w *a* ψ*-state if* $\psi \in \mathtt{L}(w)$.

In the above definition, $\mathtt{R}(w)$ marks the use of w as an intermediate state. The same maximally consistent set Δ may have different uses depending on different $\mathtt{R}(w)$. We will make use of the transitions $w \xrightarrow{\langle\psi,\bot,\chi\rangle}_c v \xrightarrow{\langle\chi^\psi,\varphi\rangle}_c w'$ where $\mathtt{R}(v) = \chi^\psi$. Note that if $\mathtt{R}(w) = \chi^\psi$ then $w \xrightarrow{\langle\chi^\psi,\varphi\rangle}_c v$ for each φ-state v. The non-trivial part of the later proof of the truth lemma is to show that adding such transitions and making them to be composed arbitrarily will not cause some $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \notin \mathtt{L}(w)$ to hold at w.

We first show that each $\Delta \in \Phi_\Gamma$ appears as $\mathtt{L}(w)$ for some $w \in S^c$.

Proposition 3.3.9 *For each* $\Delta \in \Phi_\Gamma$*, there exists* $w \in S^c$ *such that* $\mathtt{L}(w) = \Delta$.

PROOF Since $\vdash \top \to \top$, it follows by NECU that $\vdash \mathcal{U}(\top \to \top)$. Thus, we have $\mathcal{U}(\top \to \top) \in \Gamma$. It follows by EMPKhm that $\mathcal{K}h^{\mathtt{m}}(\top,\bot,\top) \in \Gamma$. It follows that $a = \langle\top,\bot,\top\rangle \in Act_\Gamma$. Since $\top \in \Delta$, it follows that $(\Delta, \top^\top) \in S^c$. □
Since $\Gamma \in \Phi_\Gamma$, it follows by Proposition 3.3.9 that $S^c \neq \emptyset$.

Proposition 3.3.5 helps us to prove the following two handy propositions which will play crucial roles in the completeness proof. Note that according to Proposition 3.3.5, to obtain that $\mathcal{U}\varphi$ in all the $\Delta \in \Phi_\Gamma$, we just need to show that φ is in all the $\Delta \in \Phi_\Gamma$, not necessarily in all the $w \in S^c$.

Proposition 3.3.10 *Given* $a = \langle\psi',\bot,\varphi'\rangle \in Act_\Gamma$*, If for each* ψ*-state* $w \in S^c$ *we have that* a *is executable at* w*, then* $\mathcal{U}(\psi \to \psi') \in \Gamma$.

PROOF Suppose that every ψ-state has an outgoing a-transition, then by the definition of R^c, ψ' is in all the ψ-states. For each $\Delta \in \Phi_\Gamma$, either $\psi \notin \Delta$, or $\psi \in \Delta$ thus $\psi' \in \Delta$. Now by the fact that Δ is maximally consistent it is not hard to show $\psi \to \psi' \in \Delta$ in both cases. By Proposition 3.3.5, $\mathcal{U}(\psi \to \psi') \in \Delta$ for all $\Delta \in \Phi_\Gamma$. It follows by $\Gamma \in \Phi_\Gamma$ that $\mathcal{U}(\psi \to \psi') \in \Gamma$. □

Proposition 3.3.11 *Given* $w \in S^c$ *and* $a = \langle\psi,\bot,\varphi'\rangle$ *or* $\langle\chi^\psi,\varphi'\rangle \in Act_\Gamma$ *such that* a *is executable at* w*, if* $\varphi \in \mathtt{L}(w')$ *for each* w' *with* $w \xrightarrow{a} w'$ *then* $\mathcal{U}(\varphi' \to \varphi) \in \Gamma$.

PROOF Firstly, we focus on the case of $a = \langle\psi,\bot,\varphi'\rangle$. For each $\Delta \in \Phi_\Gamma$ with $\varphi' \in \Delta$, we have $v = (\Delta, \varphi'^\psi) \in S^c$. Since $\langle\psi,\bot,\varphi'\rangle$ is executable at w, it means that $\psi \in \mathtt{L}(w)$. By the definition, it follows that $w \xrightarrow{a} v$. Since $\varphi \in \mathtt{L}(w')$ for each w' with $w \xrightarrow{a} w'$, it follows that $\varphi \in \mathtt{L}(v)$. Therefore, we have $\varphi \in \Delta$ for each $\Delta \in \Phi_\Gamma$ with $\varphi' \in \Delta$, namely $\varphi' \to \varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. It follows by Proposition 3.3.5 that $\mathcal{U}(\varphi' \to \varphi) \in \Gamma$.

Secondly, we focus on the case of $a = \langle\chi^\psi,\varphi'\rangle$. For each $\Delta \in \Phi_\Gamma$ with $\varphi' \in \Delta$, it follows by Proposition 3.3.9 that there exists $v \in S^c$ such that $\mathtt{L}(v) = \Delta$. Since a is executable at w, it follows that $w \xrightarrow{a} v$. Since $\varphi \in \mathtt{L}(w')$ for each w' with $w \xrightarrow{a} w'$, it follows that $\varphi \in \mathtt{L}(v)$. Therefore, we have shown that $\varphi' \in \Delta$ implies $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. It follows by Proposition 3.3.5 that $\mathcal{U}(\varphi' \to \varphi) \in \Gamma$. $\Box$

To make the proof of the truth lemma shorter, we need the following proposition.

Proposition 3.3.12 *Given a non-empty sequence $\sigma = a_1 \cdots a_n \in Act_\Gamma^*$ where $a_i = \langle \psi_i, \bot, \varphi_i\rangle$ or $a_i = \langle \chi_i^{\psi_i}, \varphi_i\rangle$ for each $1 \leq i \leq n$, we have $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_i) \in \Gamma$ for all $1 \leq i \leq n$ if for each ψ-state $w \in S^c$:*

- *σ is strongly executable at w;*
- *$w \xrightarrow{\sigma_j} t'$ implies $\chi \in \mathtt{L}(t')$ for all $1 \leq j < n$.*

PROOF If there is no ψ-state in S^c, it follows that $\neg\psi \in \mathtt{L}(w')$ for each $w' \in S^c$. It follows by Proposition 3.3.9 that $\neg\psi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. By Proposition 3.3.5, we have $\mathcal{U}\neg\psi \in \Gamma$. By UNIV, $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \bot) \in \Gamma$. Since $\vdash \bot \to \chi$ and $\vdash \bot \to \varphi$. Then by NECU, we have $\vdash \mathcal{U}(\bot \to \chi)$ and $\vdash \mathcal{U}(\bot \to \varphi)$. By UMKhm and URKhm, it is obvious that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Gamma$.

Next, assuming $v \in S^c$ is a ψ-state, we will show $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Gamma$. There are two cases: $n = 1$ or $n \geq 2$. For the case of $n = 1$, we will prove it directly; for the case of $n \geq 2$, we will prove it by induction on i.

- $n = 1$. If a_1 is of the form $\langle \chi_1^{\psi_1}, \varphi_1\rangle$, by the definition of $\xrightarrow{\langle \chi_1^{\psi_1}, \varphi_1\rangle}$ it follows that $\mathtt{R}(w) = \chi_1^{\psi_1}$ for each ψ-state w. Let χ_0 be a formula satisfying that $\vdash \chi_0 \leftrightarrow \chi_1$ and $\chi_0 \neq \chi_1$. By the rule of Replacement of Equals RE, it follows that $\langle \chi_0^{\psi_1}, \varphi_1\rangle \in Act_\Gamma$. Let $w' = (\mathtt{L}(v), \chi_0^{\psi_1})$ then it follows that $w' \in S^c$. Since $\psi \in \mathtt{L}(v)$, then we have $\psi \in \mathtt{L}(w')$. However, since $\mathtt{R}(w') = \chi_1^{\psi_1} \neq \chi_0^{\psi_1}$, $\sigma = \langle \chi_1^{\psi_1}, \varphi_1\rangle$ is not executable at the ψ-state w'. This is in contradiction with the assumption that σ is strongly executable at all ψ-states. Therefore, we know that a_1 cannot be of the form $\langle \chi_1^{\psi_1}, \varphi_1\rangle$.

  If $a_1 = \langle \psi_1, \bot, \varphi_1\rangle$, it follows that $\mathcal{K}h^{\mathtt{m}}(\psi_1, \bot, \varphi_1) \in \Gamma$. Since a_1 is executable at each ψ-state, it follows by Proposition 3.3.10 that $\mathcal{U}(\psi \to \psi_1) \in \Gamma$. Since $\mathcal{K}h^{\mathtt{m}}(\psi_1, \bot, \varphi_1) \in \Gamma$, it follows by ULKhm that $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi_1) \in \Gamma$. By NECU and UMKhm, it is clear that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_1) \in \Gamma$.

- $n \geq 2$. By induction on i, next we will show that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_i) \in \Gamma$ for each $1 \leq i \leq n$. For the case of $i = 1$, with the similar proof as in the case of $n = 1$, we can show that a_1 can only be $\langle \psi_1, \bot, \varphi_1\rangle$ and $\mathcal{U}(\psi \to \psi_1) \in \Gamma$. Therefore by UKhm we have $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_1) \in \Gamma$. Under the induction hypothesis (IH) that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_i) \in \Gamma$ for each $1 \leq i \leq k$, we will show that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_{k+1}) \in \Gamma$, where $1 \leq k \leq n-1$. Because σ is strongly executable at v, it follows that there are $w', v' \in S^c$ such that

  $$v \xrightarrow{a_1} \cdots \xrightarrow{a_{k-1}} w' \xrightarrow{a_k} v' \xrightarrow{a_{k+1}} \cdots \xrightarrow{a_n} t.$$

  Moreover, for each t' with $w' \xrightarrow{a_k} t'$ we have $\chi \in \mathtt{L}(t')$. It follows by Proposition 3.3.11 that $\mathcal{U}(\varphi_k \to \chi) \in \Gamma$ ($\blacktriangle$). Proceeding, there are two cases of a_{k+1}:

– $a_{k+1} = \langle \psi_{k+1}, \bot, \varphi_{k+1} \rangle$. Since σ is strongly executable at v, it follows that for each t' with $w' \xrightarrow{a_k} t'$ we know that a_{k+1} is executable at each t'. It follows by the definition of $\xrightarrow{\langle \psi_{k+1}, \bot, \varphi_{k+1} \rangle}$ that $\psi_{k+1} \in \mathtt{L}(t')$. Moreover, since a_k is executable at w', it follows by Proposition 3.3.11 that $\mathcal{U}(\varphi_k \to \psi_{k+1}) \in \Gamma$. Since $a_{k+1} \in Act_\Gamma$, it then follows that $\mathcal{K}h^{\mathtt{m}}(\psi_{k+1}, \bot, \varphi_{k+1}) \in \Gamma$. It follows by ULKhm that $\mathcal{K}h^{\mathtt{m}}(\varphi_k, \bot, \varphi_{k+1}) \in \Gamma$. Since $\vdash \mathcal{U}(\bot \to \chi)$, it follows by UMKhm that $\mathcal{K}h^{\mathtt{m}}(\varphi_k, \chi, \varphi_{k+1}) \in \Gamma$. Since by IH we have that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_k) \in \Gamma$, It follows from (▲) and COMPKhm that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_{k+1}) \in \Gamma$.

– $a_{k+1} = \langle \chi_{k+1}^{\psi_{k+1}}, \varphi_{k+1} \rangle$. Since σ is strongly executable at v, it follows that for each t' with $w' \xrightarrow{a_k} t'$ we know that a_{k+1} is executable at t'. Then we have that $\mathtt{R}(t') = \chi_{k+1}^{\psi_{k+1}}$ for each t' with $w' \xrightarrow{a_k} t'$.

  Note that the action a_k cannot be of the form $\langle \chi_k^{\psi_k}, \varphi_k \rangle$. Suppose it can be, let $v'' = (\mathtt{L}(v'), \chi_0^{\psi_{k+1}})$ where $\vdash \chi_0 \leftrightarrow \chi_{k+1}$ and $\chi_0 \neq \chi_{k+1}$. Since $w' \xrightarrow{a_k} v'$, it follows that $\varphi_k \in \mathtt{L}(v')$. Then it follows by the definition of transitions that $w' \xrightarrow{a_k} v''$. However, we know that $\mathtt{R}(v'') \neq \chi_{k+1}^{\psi_{k+1}}$ thus $a_{k+1} = \langle \chi_{k+1}^{\psi_{k+1}}, \varphi_{k+1} \rangle$ is not executable at v''. This is in contradiction with the strong executability. Therefore, we know that a_k cannot be of the form $\langle \chi_k^{\psi_k}, \varphi_k \rangle$.

  Now $a_k = \langle \psi_k, \bot, \varphi_k \rangle$. Since $w' \xrightarrow{a_k} v'$ and $a_{k+1} = \langle \chi_{k+1}^{\psi_{k+1}}, \varphi_{k+1} \rangle$ is executable at v', we have $\mathtt{R}(v') = \varphi_k^{\psi_k} = \chi_{k+1}^{\psi_{k+1}}$ by definition of transitions. It follows that $\psi_k = \psi_{k+1}$ and $\varphi_k = \chi_{k+1}$. Since $a_{k+1} \in Act_\Gamma$, it follows that $\mathcal{K}h^{\mathtt{m}}(\psi_{k+1}, \chi_{k+1}, \varphi_{k+1}) \in \Gamma$. Thus, we have $\mathcal{K}h^{\mathtt{m}}(\psi_k, \varphi_k, \varphi_{k+1}) \in \Gamma$. By (▲) and UMKhm we then have that $\mathcal{K}h^{\mathtt{m}}(\psi_k, \chi, \varphi_{k+1}) \in \Gamma$ (▼). If $k = 1$, by Proposition 3.3.10 it is easy to show that $\mathcal{U}(\psi \to \psi_1) \in \Gamma$. Then by ULKhm we have $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_{k+1}) \in \Gamma$. If $k > 1$, there is a state w'' such that

$$v \xrightarrow{a_1} \cdots \xrightarrow{a_{k-2}} w'' \xrightarrow{a_{k-1}} w' \xrightarrow{a_k} v' \xrightarrow{a_{k+1}} \cdots \xrightarrow{a_n} t.$$

  Since σ is strongly executable at v, it follows that for each t' with $w'' \xrightarrow{a_{k-1}} t'$ we have a_k is executable at t'. It follows by the definition of $\xrightarrow{\langle \psi_k, \bot, \varphi_k \rangle}$, it follows that $\psi_k \in \mathtt{L}(t')$ for each t' with $w'' \xrightarrow{a_{k-1}} t'$. Since a_{k-1} is executable at w'', it follows by Proposition 3.3.11 that $\mathcal{U}(\varphi_{k-1} \to \psi_k) \in \Gamma$.Moreover, since $v \xrightarrow{\sigma_{k-1}} t'$ for each t' with $w'' \xrightarrow{a_{k-1}} t'$, it follows that $\chi \in \mathtt{L}(t')$. Thus by Proposition 3.3.11 again, we have $\mathcal{U}(\varphi_{k-1} \to \chi) \in \Gamma$. Since we have proved (▼), it follows by ULKhm that $\mathcal{K}h^{\mathtt{m}}(\varphi_{k-1}, \chi, \varphi_{k+1}) \in \Gamma$. Since by IH we have $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_{k-1}) \in \Gamma$, it follows by COMPKhm that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi_{k+1}) \in \Gamma$.

□

Now we are ready to prove the truth lemma.

Lemma 3.3.13 (Truth Lemma) *For each φ and each $w \in S^c$, we have $\mathcal{M}^c_\Gamma, w \vDash \varphi$ iff $\varphi \in \mathtt{L}(w)$.*

PROOF Boolean cases are trivial, and we only focus on the case of $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi)$.

Left to Right: If there is no state w' such that $\mathcal{M}^c_\Gamma, w' \vDash \psi$, it follows by IH that $\neg\psi \in \mathtt{L}(w')$ for each $w' \in S^c$. It follows by Proposition 3.3.9 that $\neg\psi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. By Proposition 3.3.5, we have $\mathcal{U}\neg\psi \in \mathtt{L}(w)$. By UNIV, $\mathcal{K}h^{\mathtt{m}}(\psi,\bot,\bot) \in \mathtt{L}(w)$. Since $\vdash \bot \to \chi$ and $\vdash \bot \to \varphi$. Then by NECU, we have $\vdash \mathcal{U}(\bot \to \chi)$ and $\vdash \mathcal{U}(\bot \to \varphi)$. By UMKhm and URKhm, it is obvious that $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \in \mathtt{L}(w)$.

Next, assuming $\mathcal{M}^c_\Gamma, v \vDash \psi$ for some $v \in S^c$, we will show $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \in \mathtt{L}(w)$. Since $\mathcal{M}^c_\Gamma, w \vDash \mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi)$, it follows that there exists $\sigma \in \mathit{Act}^*$ such that for each $\mathcal{M}^c_\Gamma, w' \vDash \psi$: σ is strongly χ-executable at w' and $\mathcal{M}^c_\Gamma, v' \vDash \varphi$ for all v' with $w' \xrightarrow{\sigma} v'$. There are two cases: σ is empty or not.

- $\sigma = \epsilon$. This means that $\mathcal{M}^c_\Gamma, w' \vDash \varphi$ for each $\mathcal{M}^c_\Gamma, w' \vDash \psi$. It follows by IH that $\psi \in \mathtt{L}(w')$ implies $\varphi \in \mathtt{L}(w')$. Thus, we have $\psi \to \varphi \in \mathtt{L}(w')$ for all $w' \in S^c$. By Proposition 3.3.9, we have $\psi \to \varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. It follows by Proposition 3.3.5 that $\mathcal{U}(\psi \to \varphi) \in \mathtt{L}(w)$. It then follows by EMPKhm that $\mathcal{K}h^{\mathtt{m}}(\psi,\bot,\varphi) \in \mathtt{L}(w)$. By NECU and UMKhm, it is easy to show that $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \in \mathtt{L}(w)$.

- $\sigma = a_1 \cdots a_n$ where for each $1 \le i \le n$, $a_i = \langle \psi_i, \bot, \varphi_i \rangle$ or $a_i = \langle \chi_i^{\psi_i}, \varphi_i \rangle$. Since σ is strongly χ-executable at each w' with $\mathcal{M}^c_\Gamma, w' \vDash \psi$, it follows by IH that for each ψ-state w': σ is strongly executable at w' and $w' \xrightarrow{\sigma_j} t'$ implies $\chi \in \mathtt{L}(t')$ for all $1 \le j < n$. By Proposition 3.3.12, we have that $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi_n) \in \mathtt{L}(v)$. Since $\mathcal{M}^c_\Gamma, v \vDash \psi$ and σ is strongly χ-executable at v and $\mathcal{M}^c_\Gamma, v'' \vDash \varphi$ for each v'' with $v \xrightarrow{\sigma} v''$, it follows that there exists v' such that a_n is executable at v' and $\mathcal{M}^c_\Gamma, v'' \vDash \varphi$ for each v'' with $v' \xrightarrow{a_n} v''$. (Please note that $v' = v$ if $n = 1$.) Note that a_n is either $\langle \psi_n, \bot, \varphi_n \rangle$ or $\langle \chi_n^{\psi_n}, \varphi_n \rangle$. It follows by Proposition 3.3.11 and IH that if $\mathcal{U}(\varphi_n \to \varphi) \in \Gamma$, then we have $\mathcal{U}(\varphi_n \to \varphi) \in \mathtt{L}(v)$. It follows by URKhm and Proposition 3.3.3 that $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \in \mathtt{L}(w)$.

This completes the proof for $w \vDash \mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi)$ implies $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \in \mathtt{L}(w)$.

Right to Left: Suppose that $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \in \mathtt{L}(w)$. We need to show that $\mathcal{M}^c_\Gamma, w \vDash \mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi)$. There are two cases: there is a state $w' \in S^c$ such that $\mathcal{M}^c_\Gamma, w' \vDash \psi$ or not. If there is no such state, it follows $\mathcal{M}^c_\Gamma, w \vDash \mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi)$.

For the second case, let w' be a state such that $\mathcal{M}^c_\Gamma, w' \vDash \psi$. It follows by IH that $\psi \in \mathtt{L}(w')$. Since we already have $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \in \mathtt{L}(w)$, it follows by Proposition 3.3.3 that $\mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi) \in \Gamma$. Since $\vdash \mathcal{U}(\chi \to \top)$, it follows by UMKhm that $\mathcal{K}h^{\mathtt{m}}(\psi,\top,\varphi) \in \Gamma$. It follows by Proposition 3.3.6 that there exists $\Delta' \in \Phi_\Gamma$ such that $\varphi \in \Delta'$. There are two cases: $\mathcal{K}h^{\mathtt{m}}(\psi,\bot,\varphi) \in \Gamma$ or not.

- $\mathcal{K}h^{\mathtt{m}}(\psi,\bot,\varphi) \in \Gamma$. It follows that $a = \langle \psi, \bot, \varphi \rangle \in \mathit{Act}_\Gamma$. Therefore, we have $v = (\Delta', \varphi^\psi) \in S^c$. Since $\psi \in \mathtt{L}(w')$, it follows that $w' \xrightarrow{a} v$. Thus, a is strongly χ-executable at w'. What is more, $\varphi \in \mathtt{L}(v')$ for each v' with $w' \xrightarrow{a} v'$ by the definition of the transition. It follows by IH that $\mathcal{M}^c_\Gamma, v' \vDash \varphi$ for all v' with $w' \xrightarrow{a} v'$. Therefore, we have $\mathcal{M}^c_\Gamma, w \vDash \mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi)$ witnessed by a single step σ.

- $\neg\mathcal{K}h^{\mathtt{m}}(\psi,\bot,\varphi) \in \Gamma$. It follows by ONEKhm that $\mathcal{K}h^{\mathtt{m}}(\psi,\bot,\chi) \in \Gamma$. We then have $a = \langle \psi, \bot, \chi \rangle \in \mathit{Act}_\Gamma$ and $b = \langle \chi^\psi, \varphi \rangle \in \mathit{Act}_\Gamma$. Since $\mathcal{K}h^{\mathtt{m}}(\psi,\bot,\chi) \in \Gamma$ and $\vdash \mathcal{U}(\bot \to \top)$, it follows by UMKhm that $\mathcal{K}h^{\mathtt{m}}(\psi,\top,\chi) \in \Gamma$. It follows by

Proposition 3.3.6 that there exists $\Delta'' \in \Phi_\Gamma$ such that $\chi \in \Delta''$. Therefore, we have $t = (\Delta'', \chi^\psi) \in S^c$. Since there exists $\Delta' \in \Phi_\Gamma$ with $\varphi \in \Delta'$, it follows by Proposition 3.3.5 that there is $t' \in S^c$ such that $\mathtt{L}(t') = \Delta'$. Now, starting with any ψ-state, a is clearly executable and it will lead to a χ-state, and then by a b step we will reach all the φ states. Therefore, by IH, we have that ab is strongly χ-executable at w', and that for all v' with $w' \xrightarrow{ab} v'$ we have $\mathcal{M}^c_\Gamma, v' \vDash \varphi$. Therefore, we have $\mathcal{M}^c_\Gamma, w \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$. Note that we do need a 2-step σ in this case.

□

Now due to Proposition 3.3.4, each $\mathbb{SKHM}$-consistent set of formulas can be extended to a maximally consistent set Γ. Due to the Truth Lemma 3.3.13, we have $\mathcal{M}^c_\Gamma, (\Gamma, \top^\top) \vDash \Gamma$. The completeness of $\mathbb{SKHM}$ follows immediately.

Theorem 3.3.14 (Completeness) $\mathbb{SKHM}$ *is strongly complete w.r.t. the class of all models.*

3.4 Decidability

In this section, we will show that the problem whether a formula φ_0 is satisfiable is decidable. The idea is that we will show that φ_0 has a bounded small model if φ_0 is satisfiable. From the truth lemma and the completeness theorem, we already know that if φ_0 is satisfiable then it is satisfiable in a canonical model. However, a canonical model is infinite. Our method is to make the filtration of a canonical model through a finite set generated by φ_0 and to show the filtration model is a bounded small model.

Firstly, we will define the set through which we will make the filtration of a canonical model. Different from the standard filtration method in modal logic (see Blackburn et al. (2001)), this set is not only closed under subformulas, but we add $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ and $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \chi)$ for each formula $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ in the set.

Definition 3.4.1 (Subformula Closed) *A set of formulas Δ is* closed under subformulas *if for all formulas φ, ψ, χ: if $\neg\varphi \in \Delta$ then $\varphi \in \Delta$; if $\varphi \wedge \psi \in \Delta$ then $\varphi \in \Delta$ and $\psi \in \Delta$; if $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Delta$ then $\psi \in \Delta$, $\chi \in \Delta$ and $\varphi \in \Delta$.*

Definition 3.4.2 *Let Φ be a subformula closed set. $cl(\Phi)$ is the smallest set such that:*

- $\Phi \subseteq cl(\Phi)$;
- *if $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Phi$ and $\chi \neq \bot$ then $\pm\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi), \pm\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \chi) \in cl(\Phi)$.*

If Φ is a set generated by a single formula, it is obvious that $cl(\Phi)$ is finite. The formula set $cl(\Phi)$ is the set we need for our filtration method.

Before building the filtration model, we firstly define the action set that is to be part of the filtration model. From the proof of the truth lemma, we know that each formula $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ that is true in a canonical model is witnessed by the plan $\langle \psi, \bot, \varphi \rangle$ or the plan $\langle \psi, \bot, \chi \rangle \langle \chi^\psi, \varphi \rangle$. Since in the filtration model we only care about the formulas in the closure $cl(\Phi)$, thus we only need the actions in $Act_{\Gamma\cap cl(\Phi)}$. Remember Definition 3.3.7:

$$\begin{aligned} Act_{\Gamma\cap cl(\Phi)} =& \{\langle\psi,\bot,\varphi\rangle \mid \mathcal{K}h^{\mathtt{m}}(\psi,\bot,\varphi)\in(\Gamma\cap cl(\Phi))\}\cup \\ & \{\langle\chi^{\psi},\varphi\rangle \mid \mathcal{K}h^{\mathtt{m}}(\psi,\chi,\varphi),\neg\mathcal{K}h^{\mathtt{m}}(\psi,\bot,\varphi)\in(\Gamma\cap cl(\Phi))\} \end{aligned}$$

It is obvious that $Act_{\Gamma\cap cl(\Phi)}\subseteq Act_{\Gamma}$.

Let Γ be a maximally consistent set and let $\mathcal{M}^c_\Gamma$ be the model defined in Definition 3.3.8. Next we will define the filtration of $\mathcal{M}^c_\Gamma$ through $cl(\Phi)$. Before that, we first define an equivalence relation $\leftrightsquigarrow_{cl(\Phi)}$ over the state set S^c. Let s,t be two states in S^c. The equivalence relation $\leftrightsquigarrow_{cl(\Phi)}$ is defined as below.

$$\begin{aligned} s\leftrightsquigarrow_{cl(\Phi)} t \iff & \mathcal{M}^c_\Gamma,s\vDash\varphi \text{ iff } \mathcal{M}^c_\Gamma,t\vDash\varphi \text{ for all } \varphi\in cl(\Phi)\text{, and} \\ & a \text{ is executable on } s \text{ iff } a \text{ is executable on } t \text{ for all } a\in Act_{\Gamma\cap cl(\Phi)}. \end{aligned}$$

We denote the equivalence class of a state s of $\mathcal{M}^c_\Gamma$ with respect to $\leftrightsquigarrow_{cl(\Phi)}$ by $|s|_{cl(\Phi)}$, or simply $|s|$ if no confusion will arise.

Definition 3.4.3 (Filtration) *The filtration of $\mathcal{M}^c_\Gamma$ through $cl(\Phi)$, denoted by $(\mathcal{M}^c_\Gamma)^f=\langle S^f,Act^f,R^f,V^f\rangle$, is defined as below.*

- $S^f=\{|s|_{cl(\Phi)}\mid s\in S^c\}$.
- $Act^f=Act_{\Gamma\cap cl(\Phi)}$.
- *For all $a\in Act^f$, $|s|\xrightarrow{a}|t|$ iff there are $s'\in|s|$ and $t'\in|t|$ such that $s'\xrightarrow{a}t'$.*
- *For all $p\in cl(\Phi)$, $p\in V^f(|s|)$ iff $\mathcal{M}^c_\Gamma,s\vDash p$.*

To show the filtration lemma, the key is to show that a formula has a strongly executable plan in the original canonical model if and only if it has a strongly executable plan in the filtration model, which is what the following two propositions show.

Proposition 3.4.4 *Given $\sigma=a_1\cdots a_n\in(Act^f)^*$, $n\in\mathbb{N}$ and $|s|\in S^f$, if σ is strongly executable on $|s|$ in $(\mathcal{M}^c_\Gamma)^f$ then σ is strongly executable on all $s'\in|s|$ in $\mathcal{M}^c_\Gamma$.*

PROOF We prove it by induction on the length of σ. It is obvious if $\sigma=\epsilon$. Next we will show that the proposition holds for $\sigma=a_1\cdots a_{n+1}$.

If $a_1\cdots a_{n+1}$ is strongly executable on $|s|$, it is obvious that $a_1\cdots a_n$ is strongly executable on $|s|$. It follows by the IH that $a_1\cdots a_n$ is strongly executable on all $s'\in|s|$. Given $s'\in|s|$, we will show that $a_1\cdots a_{n+1}$ is strongly executable on s'. Since $a_1\cdots a_n$ is strongly executable on $|s|$, by Definition 3.2.3, we only need to show that if $s\xrightarrow{\sigma_n}t$ then a_{n+1} is executable on t.

If $s\xrightarrow{\sigma_n}t$, that is, $s\xrightarrow{a_1}s_1\cdots\xrightarrow{a_n}t$, it follows that $|s|\xrightarrow{\sigma_n}|t|$. Since $a_1\cdots a_{n+1}$ is strongly executable on $|s|$, it follows that a_{n+1} is executable on $|t|$. Therefore, there exists $t'\in|t|$ such that a_{n+1} is executable on t'. It follows by the definition of $\leftrightsquigarrow$ that a_{n+1} is executable on t. $\square$

Proposition 3.4.5 *Given $a = \langle \psi, \bot, \chi \rangle \in Act^f$, $b = \langle \chi^\psi, \varphi \rangle \in Act^f$, and $s \in S^c$, the following two propositions hold.*

1. *If ab is strongly executable on s then ab is strongly executable on $|s|$.*

2. *If $|s| \xrightarrow{a} |t|$ and $|t| \xrightarrow{b} |v|$ for some $|t|, |v| \in S^f$ then there exits $t' \in |t|$ and $v' \in |v|$ such that $s \xrightarrow{a} t'$ and $t' \xrightarrow{b} v'$.*

PROOF In the following, we will prove the two propositions, respectively.

1. If ab is strongly executable at s, it follows that there are $t, v \in S^c$ such that $s \xrightarrow{a} t$ and $t \xrightarrow{b} v$. What is more, by the form of the actions a and b, it follows by Definition 3.3.8 that $t = (\Delta, \chi^\psi)$, $\psi \in \mathtt{L}(s)$, and $\varphi \in \mathtt{L}(v)$.

   Since a is executable on s, it follows that a is also executable on $|s|$. To show that ab is strongly executable on $|s|$, we only need to show that b is executable on each $|t_0| \in S^f$ with $|s| \xrightarrow{a} |t_0|$. If $|s| \xrightarrow{a} |t_0|$, it follows that there are $s' \in |s|$ and $t' \in |t_0|$ such that $s' \xrightarrow{a} t'$. By the form of a, it follows that $t' = (\Delta', \chi^\psi)$. Since $\varphi \in \mathtt{L}(v)$, it follows by Definition 3.3.8 that $t' \xrightarrow{b} v$. Therefore, we have $|t'| \xrightarrow{b} |v|$, that is, $|t_0| \xrightarrow{b} |v|$. Thus, b is executable on $|t_0|$.

2. If $|s| \xrightarrow{a} |t|$, it follows that a is executable on some $s' \in |s|$. By Definition 3.3.8 and the form of the action a, it follows that $\psi \in \mathtt{L}(s')$ because a is executable on s'. It follows by the truth lemma that $\mathcal{M}^c_\Gamma, s' \vDash \psi$. Since $s' \in |s|$, it follows that $\mathcal{M}^c_\Gamma, s \vDash \psi$. By the truth lemma again, we have $\psi \in \mathtt{L}(s)$.

   If $|t| \xrightarrow{b} |v|$, it follows that there are $t' \in |t|$ and $v' \in |v|$ such that $t' \xrightarrow{b} v'$. By Definition 3.3.8 and the form of the action b, it follows that $t' = (\Delta', \chi^\psi)$. Since $\psi \in \mathtt{L}(s)$, it follows that $s \xrightarrow{a} t'$. Thus, we have $s \xrightarrow{a} t'$ and $t' \xrightarrow{b} v'$.

□

Now we are ready to prove the filtration lemma.

Lemma 3.4.6 (Filtration Lemma) *For each formula $\varphi \in cl(\Phi)$, $(\mathcal{M}^c_\Gamma)^f, |s_0| \vDash \varphi$ iff $\mathcal{M}^c_\Gamma, s_0 \vDash \varphi$.*

PROOF We prove it by induction on φ. We restrict our attention to the case of $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$; the other cases are trivial.

Left-to-Right. If $(\mathcal{M}^c_\Gamma)^f, |s_0| \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$, next we will show that $\mathcal{M}^c_\Gamma, s_0 \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$. Firstly it follows that there exists $\sigma \in Act^f$ such that for all $|s'|$ with $(\mathcal{M}^c_\Gamma)^f, |s'| \vDash \psi$: σ is strongly χ-executable on $|s'|$ and $(\mathcal{M}^c_\Gamma)^f, |t'| \vDash \varphi$ for all $|t'|$ with $|s'| \xrightarrow{\sigma} |t'|$.

To show $\mathcal{M}^c_\Gamma, s_0 \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$, let s be a state such that $\mathcal{M}^c_\Gamma, s \vDash \psi$. Then we only need to show that σ is strongly χ-executable on s and $\mathcal{M}^c_\Gamma, t \vDash \varphi$ for all t with $s \xrightarrow{\sigma} t$. Since $\mathcal{M}^c_\Gamma, s \vDash \psi$, it follows by the IH that $(\mathcal{M}^c_\Gamma)^f, |s| \vDash \psi$. Therefore, σ is strongly executable on $|s|$. It follows by Proposition 3.4.4 that σ is strongly executable on s. Let $\sigma = a_1 \cdots a_n$, then for each $0 \leq k \leq n$: $s \xrightarrow{\sigma_k} t$ implies $|s| \xrightarrow{\sigma_k} |t|$. It follows by the IH that σ is strongly χ-executable on s and $\mathcal{M}^c_\Gamma, t \vDash \varphi$ for all t with $s \xrightarrow{\sigma} t$.

Right-to-Left. If $\mathcal{M}^c_\Gamma, s_0 \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$, next we will show that $(\mathcal{M}^c_\Gamma)^f, |s_0| \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$. If $\mathcal{M}^c_\Gamma, s_0 \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$, it follows by the truth lemma that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in S_0$, and then $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi) \in \Gamma$. There are two cases.

- $\chi = \bot$. We have $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi) \in \Gamma$, and then $\langle \psi, \bot, \varphi \rangle \in Act_\Gamma$. Since $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi) \in s_0$, by the proof of the truth lemma, it follows that for all s with $\mathcal{M}^c_\Gamma, s \vDash \psi$: $\langle \psi, \bot, \varphi \rangle$ is strongly executable on s and $\mathcal{M}^c_\Gamma, t \vDash \varphi$ for all t with $s \xrightarrow{\psi, \bot, \varphi} t$.

  Since we also have $\langle \psi, \bot, \varphi \rangle \in Act^f$, to show $(\mathcal{M}^c_\Gamma)^f, |s_0| \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$, we only need to show that for all $|s|$ with $(\mathcal{M}^c_\Gamma)^f, |s| \vDash \psi$: $\langle \psi, \bot, \varphi \rangle$ is strongly executable on $|s|$ and $(\mathcal{M}^c_\Gamma)^f, |t| \vDash \varphi$ for all $|t|$ with $|s| \xrightarrow{\psi, \bot, \varphi} |t|$. If $(\mathcal{M}^c_\Gamma)^f, |s| \vDash \psi$, it follows by the IH that $\mathcal{M}^c_\Gamma, s \vDash \psi$. Therefore, $\langle \psi, \bot, \varphi \rangle$ is (strongly) executable on s, and thus it is (strongly) executable on $|s|$. If $|s| \xrightarrow{\psi, \bot, \varphi} |t|$, it follows that there are $s' \in |s|$ and $t' \in |t|$ such that $s' \xrightarrow{\psi, \bot, \varphi} t'$. By the IH, we know that $\mathcal{M}^c_\Gamma, s' \vDash \psi$. Therefore, we have $\mathcal{M}^c_\Gamma, t' \vDash \varphi$, and it follows by the IH again that $(\mathcal{M}^c_\Gamma)^f, |t| \vDash \varphi$.

- $\chi \neq \bot$. If we still have $\mathcal{M}^c_\Gamma, s_0 \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$, by the same proof, we will have $(\mathcal{M}^c_\Gamma)^f, |s_0| \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$, and thus $(\mathcal{M}^c_\Gamma)^f, |s_0| \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$. Next we will focus on the situation of $\mathcal{M}^c_\Gamma, s_0 \nvDash \mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$. If $\mathcal{M}^c_\Gamma, s_0 \nvDash \mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$, it follows by the truth lemma that $\neg\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi) \in \Gamma$. By Definition 3.4.2, we know that $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ and $\neg\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ are also members of $cl(\Phi)$. Thus we have $b = \langle \chi^\psi, \varphi \rangle \in Act^f$. What is more, since $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ and $\neg\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \varphi)$ are in Γ, it follows by Axiom `ONEKhm` that $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \chi) \in \Gamma$. Since we also have $\mathcal{K}h^{\mathtt{m}}(\psi, \bot, \chi) \in cl(\Phi)$ by Definition 3.4.2, it follows that $a = \langle \psi, \bot, \chi \rangle \in Act^f$.

  It is obvious that the actions a and b are also in Act_Γ. By the proof of the truth lemma, we know that for all s with $\mathcal{M}^c_\Gamma, s \vDash \psi$: ab is strongly χ-executable on s and $\mathcal{M}^c_\Gamma, t \vDash \varphi$ for all t with $s \xrightarrow{ab} t$. To show that $(\mathcal{M}^c_\Gamma)^f, |s_0| \vDash \mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$, next we will show that for all $|s|$ with $(\mathcal{M}^c_\Gamma)^f, |s| \vDash \psi$: ab is strongly executable on $|s|$, and $(\mathcal{M}^c_\Gamma)^f, |t| \vDash \chi$ and $(\mathcal{M}^c_\Gamma)^f, |v| \vDash \varphi$ for all $|t|$ and $|v|$ with $|s| \xrightarrow{a} |t|$ and $|t| \xrightarrow{b} |v|$. If $(\mathcal{M}^c_\Gamma)^f, |s| \vDash \psi$, it follows by the IH that $\mathcal{M}^c_\Gamma, s \vDash \psi$. Therefore, ab is strongly executable on s. It follows by Proposition 3.4.5 that ab is strongly executable on $|s|$. If $|s| \xrightarrow{a} |t| \xrightarrow{b} |v|$, it follows by Proposition 3.4.5 that there are $t' \in |t|$ and $v' \in |v|$ such that $s \xrightarrow{a} t' \xrightarrow{b} v'$. Thus we have $\mathcal{M}^c_\Gamma, t' \vDash \chi$ and $\mathcal{M}^c_\Gamma, v' \vDash \varphi$. It follows by the IH that $(\mathcal{M}^c_\Gamma)^f, |t| \vDash \chi$ and $(\mathcal{M}^c_\Gamma)^f, |v| \vDash \varphi$.

□

Proposition 3.4.7 (Small Model Property) *If φ_0 is satisfiable then it is satisfiable in a model with at most 2^{2k} states where $k = |cl(\Phi)|$ and Φ is the subformula closure generated by φ_0.*

PROOF If φ_0 is satisfiable, it follows by soundness that it is consistent. By Lindenbaum's lemma, φ_0 can be extended to be a maximally consistent set Γ. By the truth lemma, we have $\mathcal{M}^c_\Gamma, s \vDash \varphi_0$ where $s = (\Gamma, \top^\top)$. It follows by the filtration lemma that φ_0 is satisfiable in the model $(\mathcal{M}^c_\Gamma)^f$. Each state $|t|$ in $(\mathcal{M}^c_\Gamma)^f$ corresponds to a pair (Δ, X) where $\Delta \subseteq cl(\Phi)$ is the set of formulas that are true in $|t|$ and $X \subseteq Act^f$ is the set of actions that are executable on $|t|$. Since there are at most k actions in σ^f, there are at most 2^k action set X. Since there are also at most 2^k formula set Δ, there are at most 2^{2k} pairs of (Δ, X). Therefore, there are at most 2^{2k} states in the filtration model $(\mathcal{M}^c_\Gamma)^f$. □

Theorem 3.4.8 (Decidability) *KHM is decidable.*

PROOF With the small model property, this can be proved by a standard argument presented in Blackburn et al. (2001). □

3.5 Conclusion

In this chapter, we generalized the knowing how logic presented in Wang (2015a) and proposed a ternary modal operator $\mathcal{K}h^{\mathtt{m}}(\psi, \chi, \varphi)$ to express that the agent knows how to achieve φ given ψ while taking a route that satisfies χ. We also presented a sound and complete axiomatization of this logic. Compared to the completeness proof in Wang (2015a), the proof here is more complicated. The essential difference is that, to handle the intermediate constraints, a state of the canonical model here is a pair consisting of a maximally consistent set and a marker of the form χ^ψ which indicates that this state has a $\langle \psi, \bot, \chi\rangle$-predecessor.

Moreover, we showed that the logic KHM is decidable via a filtration method. The filtration here is defined on the canonical model, and it cannot be extended to all models. That is because the $\mathcal{K}h^{\mathtt{m}}$-formulas have special witness plans in the canonical model, which allows us to select a subset of the whole action set. This filtration method can be applied to the logic presented in Wang (2015a) and shows that the logic with binary knowing-how operator $\mathcal{K}h$ is decidable.

One interesting topic related to this chapter's topic is relaxing the strong executability in the semantics. Intuitively, strongly executable plans may be too strong for knowledge-how in some cases. For example, if there is an action sequence σ in the model such that doing σ at a ψ-state will always make the agent *stop* on φ states, we can probably also say the agent knows how to achieve φ given ψ. For instance, I know how to start the engine in that old car given the precondition that it is in good condition. I just need to turn the key several times until it starts, and five times should suffice, at most. Please note that there are two kinds of states in which the agent might stop: either states that the agent achieves after doing σ successfully, or states on which the agent is unable to continue executing the remaining actions. In Chapter 4, we will discuss the logic of knowing how with this kind of plans.

Another interesting related topic is to consider contingent plans which involve conditions based on the knowledge of the agent. A *contingent* plan is a partial function on the agent's belief space. Such plans make more sense when the agent has the ability of observations during the execution of the plan. To consider contingent plans, we need to extend the model with an epistemic relation. We can then express knowledge-that and knowledge-how at the same time, and discuss their interactions in one unified logical framework. We discuss the logic of knowing how with contingent plans in Chapter 5.

For future research, besides the obvious questions of model theory of the logic, we can extend this logic with the public announcement operator. Intuitively, $[\theta]\varphi$ says that φ holds after the information θ is provided. The update of the new information amounts to the change of the background knowledge throughout the model, and this may affect the knowledge-how. For example, a doctor may not know how to treat a patient with the disease since he is worried that the only available medicine might cause some very bad side-effects. Let p mean that the patient has the disease, and let r mean that there are bad side-effects. Then this can be expressed as $\neg\mathcal{K}h^{\mathtt{m}}(p, \neg r, \neg p)$. Suppose a new scientific discovery shows that the side-effect is not possible under the relevant circumstance, then the doctor should know how to treat the patient, which can be expressed as $[\neg r]\mathcal{K}h^{\mathtt{m}}(p, \neg r, \neg p)$.[4]

[4]However, the announcement operator $[\varphi]$ is not reducible in $\mathcal{L}_{\text{KHM}}$ as discussed in Wang (2016).

Chapter 4

Knowing how with weak conformant plans[1]

4.1 Introduction

Epistemic logic proposed by Hintikka (1962) is a modal logic concerned with reasoning about knowledge. It formalizes propositional knowledge, the knowledge of the form "knowing that", by means of a modal formula $\mathcal{K}\varphi$ which expresses that the agent knows that φ holds. It interprets knowledge-that in terms of agents' uncertainty. The agent knows that φ at a state s if and only if he can rule out all the $\neg\varphi$ epistemic possibilities at s. Epistemic logic is widely applied in theoretical computer science, artificial intelligence, economics, and linguistics (see van Ditmarsch et al. (2015)).

However, knowledge is not only expressed by "knowing that", but also by other expressions, such as "knowing how", "knowing what", "knowing why", and so on. Among all these expressions, "knowing how" (and the knowledge-how that it expresses) is the most discussed.

In artificial intelligence, beginning from McCarthy and Hayes (1969) and McCarthy (1979), researchers started to study what it means for a computer program to "know how" to achieve a state of affairs φ in terms of its ability. In particular, Moore (1985) is highly influential on representation of and reasoning about knowledge-how and ability. Please note that Moore did not clearly distinguish between knowledge-how and ability. According to Moore, there are two possible ways to define the agent's knowledge-how:

(I) There exists an action a such that the agent knows that the performance of a will result in φ;

(II) The agent knows that there exists an action a such that the performance of a will result in φ.

Let $\varphi(a)$ express that performing a will make sure that φ. The first is a *de re* definition of knowledge-how ($\exists a : \mathcal{K}\varphi(a)$), and the second is a *de dicto* definition ($\mathcal{K}\exists a : \varphi(a)$). Moore pointed out that the first definition is too strong and the second is too weak. For

[1] This is based on the paper Li (2017).

example, if we only know that there exists a method to guarantee that φ but without knowing the identity of the method, we cannot say that we know how to make sure that φ. On the other hand, in real-life context, we might say to a friend that "I can be at your house at 8pm" without knowing in advance exactly what action we will adopt. Therefore, Moore proposed an adapted, but very complicated version of the definition. Moore's formalism has inspired a large body of work in artificial intelligence on knowledge and ability (see the surveys by Gochet (2013) and Ågotnes et al. (2015)).

In logic, the framework of Alternating-time Temporal Logic (ATL) (cf. Alur et al. (2002)) is concerned with reasoning about agents' abilities in game structures. By adding the knowledge operator to this framework, ATEL (van der Hoek and Wooldridge (2003)) can express that the agent knows that there is a strategy to enforce φ from the current state. However, the reading is still a *de dicto* reading of knowledge-how, and it is too weak to define knowledge-how. To solve this problem, researchers proposed different solutions (cf. Herzig et al. (2013); Belardinelli (2014); Herzig (2015)), such as making the strategy uniform, or specifying the explicit actions in the modality (e.g., knowing that performing abc will achieve φ).

In the above-mentioned works, knowledge-how is usually expressed in a very rich logical language involving quantifiers or various complicated modalities. However, starting from Plaza (1989), Hart et al. (1996), and van der Hoek and Lomuscio (2003), logicians attempted to formalize some knowledge-wh, such as "knowing whether", "knowing what" etc., as a whole modality, in a similar way in which epistemic logic deals with knowledge-that. The recent works Fan et al. (2014), Fan et al. (2015), Gu and Wang (2016), and Wang (2015a) are in line with this idea.

In particular, Wang (2015a) proposed a single-agent logic of knowing how, which includes the modal formula $\mathcal{K}h(\psi, \varphi)$ to express that the agent knows how to achieve φ given the precondition ψ. The models are labelled transition systems which reflect an agent's ability. Thus the models are also called *ability maps*. The formula $\mathcal{K}h(\psi, \chi)$ is interpreted in a *de re* reading of knowledge-how: there exists an action sequence (also called a plan) σ such that (1) performing σ at each ψ-state the agent will achieve a φ-state; and (2) the plan is not supposed to fail during the execution. In automated planning, such a plan is called a *conformant plan* (cf. Smith and Weld (1998); Ghallab et al. (2004)). Consider Figure 4.1 which represents a map of a floor in a building where the agent can go right(r) or up(u).[2] According to Wang's interpretation of knowledge-how, the agent here knows how to achieve q given p because there is a conformant plan ru (first moving right then moving up) for achieving q-states from p-states.

$$\begin{tikzcd}
 & s_6 & s_7 : q & s_8 : q & \\
s_1 \arrow[r, "r"] & s_2 : p \arrow[u, "u"] \arrow[ur, "u"] \arrow[r, "r"] & s_3 : p \arrow[u, "u"] \arrow[r, "r"] & s_4 \arrow[u, "u"] \arrow[r, "r"] & s_5 : q
\end{tikzcd}$$

Figure 4.1

However, the demands that a conformant plan asks may be too strong, in the sense that the execution of the plan will *never* fail. Intuitively, we are still comfortable to say

[2] This is a variant of the running example used in Wang and Li (2012).

"we know how to achieve φ given ψ" if we will always end up with a φ-state when the execution of the plan terminates, whether or not all parts of the plan have been completed. For example, let q be true only on the state s_5 in Figure 4.1. Then there will be no conformant plans for achieving the only q-state s_5 from p-states, but we still say that "we know how to achieve the q-state from p-states" because we can get there by moving right at most three times. The plan of moving right three times is not a conformant plan since the execution of the plan starting from s_3 will fail at s_5, but this plan will still guarantee our achieving the q-state s_5 in the sense that we will always end up with s_5 when the execution of the plan terminates. We call it a weak conformant plan. A *weak conformant plan* for achieving φ-states from ψ-states is a finite linear action sequence such that the execution of the action sequence at each ψ-state will always terminate on a φ-state, either successfully or not. Intuitively, a weak conformant plan is enough for our knowing how to achieve φ given ψ.

In this chapter, we interpret knowledge-how as there being a weak conformant plan for the agent achieving the goal. Compared to the interpretation of Wang (2015a), our interpretation is weaker, but more realistic. We also present a sound and complete axiomatic system. It shows that this weaker interpretation results in a weaker logic. The composition axiom in Wang (2015a)

$$(\mathcal{K}h(p, r) \wedge \mathcal{K}h(r, q)) \to \mathcal{K}h(p, q)$$

is not valid under this weaker interpretation. Even though the logic is weaker, the proof of its completeness is non-trivial. We also define an alternative nonstandard semantics. By reducing a decidable problem on our weaker semantics to a decidable problem on this alternative nonstandard semantics, we show that this logic is decidable.

The rest of the chapter is organized as follows: Section 4.2 introduces the language and semantics and presents a deductive system; Section 4.3 shows the completeness of the deductive system; Section 4.4 proposes an alternative semantics for our logic and shows that our logic is decidable; in the last section, we conclude with future directions.

4.2 The logic KHW

This section will introduce the logic of knowing how with weak conformant plans, and we denote the logic as KHW.

4.2.1 Syntax and semantics

In this section, we will introduce the language and the semantics. The language has first been proposed in Wang (2015a). Differently from the knowing-how modality $\mathcal{K}h$ used in Wang (2015a), we use the modality $\mathcal{K}h^{\mathtt{w}}$ here, and the superscript indicates that this logic is weaker.

Definition 4.2.1 (Language) *Given a countable set of proposition letters $\mathbf{P}$, the language $\mathcal{L}_{KHW}$ is defined as follows:*

$$\varphi := \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathcal{K}h^{\mathtt{w}}(\varphi, \varphi)$$

where $p \in \mathbf{P}$. We will often omit parentheses around expressions when doing so should not cause confusion. We use the standard abbreviations $\bot, \varphi \vee \psi$ and $\varphi \to \psi$. We define

$\mathcal{U}\varphi$ as $\mathcal{K}h^{\mathtt{w}}(\neg\varphi, \bot)$. $\mathcal{U}$ is intended to be a universal modality, and it will become clear after we define the semantics.

Intuitively, the formula $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi)$ expresses that the agent knows how to guarantee φ given ψ since she has a weak conformant plan to achieve φ-states from each ψ-state.

The language is interpreted on models which are labelled transition systems. The model is also called an ability map because it represents the agent's abilities, i.e. it illustrates what actions the agent can do in each state.

Definition 4.2.2 (Model) *A model (also called an ability map) is essentially a labelled transition system (S, Act, R, V) where:*

- *S is a non-empty set of states;*
- *Act is a set of actions (or labels);*
- *$R : Act \to 2^{S\times S}$ is a collection of transitions labelled by actions in Act;*
- *$V : S \to 2^{\mathbf{P}}$ is a valuation function.*

We write $s \xrightarrow{a} t$ or $t \in R_a(s)$ if $(s, t) \in R(a)$. For a sequence $\sigma = a_1 \dots a_n \in Act^*$, we write $s \xrightarrow{\sigma} t$ if there exist $s_2 \dots s_n$ such that $s \xrightarrow{a_1} s_2 \xrightarrow{a_2} \cdots \xrightarrow{a_{n-1}} s_n \xrightarrow{a_n} t$. Note that σ can be the empty sequence ϵ (when $n = 0$), and we set $s \xrightarrow{\epsilon} s$ for any s. Let σ_k be the initial segment of σ up to a_k for $k \leq |\sigma|$. In particular let $\sigma_0 = \epsilon$. We say that σ is executable *at s if there is t such that $s \xrightarrow{\sigma} t$.*

Note that the labels in Act do not appear in the language. The graph in Figure 4.1 represents a model. We also call an action sequence a plan. We say a plan σ is executable in a state s if there exists a state t such that $s \xrightarrow{\sigma} t$.

As discussed in Section 4.1, we will interpret knowledge-how as there being a weak conformant plan for achieving the goal, that is, the agent will always terminate on a goal state when performing the plan. Next, we will formally define the set of states on which the agent will terminate when performing a plan.

Definition 4.2.3 (Terminal States) *Given a state $s \in S$ and an action sequence $\sigma = a_1 \cdots a_n \in Act^*$, $\mathtt{TermiSs}(s, \sigma)$ is the set of states on which the agent might terminate if she performs σ in s. Formally, it is defined as*

$$\mathtt{TermiSs}(s, \sigma) = \{t \mid s \xrightarrow{\sigma} t, \text{ or } \exists i < n : s \xrightarrow{\sigma_i} t \text{ and } t \text{ has no } a_{i+1} \text{ successor}\}.$$

In particular, let $\mathtt{TermiSs}(s, \epsilon) = \{s\}$. If $s \xrightarrow{\sigma} t$ for all $t \in \mathtt{TermiSs}(s, \sigma)$, we say σ is strongly executable *at s.*

Considering the following model, we have $\mathtt{TermiSs}(s_1, ab) = \{s_3, s_4\}$.

$$\begin{tikzcd}[row sep=small] & s_2 \arrow[r, "b"] & s_4 : q \\ s_1 : p \arrow[ur, "a"] \arrow[dr, "a"'] & & \\ & s_3 & \end{tikzcd}$$

Definition 4.2.4 (Semantics) *Suppose s is a state in a model $\mathcal{M} = (S, Act, R, V)$. Then we inductively define the notion of a formula φ being satisfied (or true) in $\mathcal{M}$ at state s as follows:*

$$\begin{array}{lcl}
\mathcal{M}, s \vDash \top & & \textit{always} \\
\mathcal{M}, s \vDash p & \iff & p \in V(s). \\
\mathcal{M}, s \vDash \neg\varphi & \iff & \mathcal{M}, s \nvDash \varphi. \\
\mathcal{M}, s \vDash (\varphi \wedge \psi) & \iff & \mathcal{M}, s \vDash \varphi \textit{ and } \mathcal{M}, s \vDash \psi. \\
\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{w}}(\psi, \varphi) & \iff & \textit{there exists } \sigma \in Act^* \textit{ such that for each } w \in [\![\psi]\!] \\
 & & \textit{and each } t \in \mathtt{TermiSs}(w, \sigma) \textit{ we have } \mathcal{M}, t \vDash \varphi.
\end{array}$$

where $[\![\psi]\!] = \{s \in S \mid \mathcal{M}, s \vDash \psi\}$.

We also call the semantics defined here as the *standard semantics*, to distinguish it from the nonstandard semantics defined in Section 4.4. Now we can also check that the operator $\mathcal{U}$ defined by $\mathcal{K}h^{\mathtt{w}}(\neg\varphi, \bot)$ is indeed a *universal modality*:

$$\boxed{\mathcal{M}, s \vDash \mathcal{U}\varphi \quad \Leftrightarrow \quad \text{for all } t \in S, \mathcal{M}, t \vDash \varphi}$$

Under this semantics, the composition axiom in Wang (2015a),

$$(\mathcal{K}h^{\mathtt{w}}(p, r) \wedge \mathcal{K}h^{\mathtt{w}}(r, q)) \to \mathcal{K}h^{\mathtt{w}}(p, q),$$

is not valid. The following example presents a model in which the composition axiom is not true.

Example 4.2.5 *Model $\mathcal{M}$ is depicted as follows.*

$$s_1 : p \xrightarrow{a} s_3 : r \xrightarrow{b} s_5 : q$$

$$s_2 : p, r \xrightarrow{b} s_4 : q$$

- $\mathcal{M}, s_1 \vDash \mathcal{K}h^{\mathtt{w}}(p, r)$ *since there is a weak conformant plan a. Please note that after executing a on s_2 the agent will terminate on s_2 itself.*
- $\mathcal{M}, s_1 \vDash \mathcal{K}h^{\mathtt{w}}(r, q)$ *since there is a weak conformant plan b. After executing b on each r-states, either s_3 or s_2, the agent will achieving on a q-state.*
- $\mathcal{M}, s_1 \nvDash \mathcal{K}h^{\mathtt{w}}(p, q)$ *since there are no weak conformant plans for achieving q-states from p-states. Particularly, ab is not a weak conformant plan. The performance of ab on s_1 will result in a q-state s_5, but executing ab on p-state s_2 will terminate on s_2 itself.*

The composition of two weak conformant plans might not be a weak conformant plan any more. Just as shown in Example 4.2.5, a is a weak conformant plan for achieving r-states from p-states, and b is a weak conformant plan for achieving q-states from r-states, but the composition ab is not a weak conformant plan for achieving q-states from p-states. There is no weak conformant plan for achieving q-states from p-states in this example.

4.2.2 A deductive system

In this subsection, we provide a Hilbert-style proof system for the logic KHW and show it is sound on the standard semantics.

Definition 4.2.6 ($\mathbb{SWKH}$ System) *The axiomatic system $\mathbb{SWKH}$ is shown in Table 4.1. We write $\mathbb{SWKH} \vdash \varphi$ (or sometimes just $\vdash \varphi$) to mean that the formula φ is derivable in the axiomatic system $\mathbb{SWKH}$; the negation of $\mathbb{SWKH} \vdash \varphi$ is written $\mathbb{SWKH} \nvdash \varphi$ (or just $\nvdash \varphi$). To say that a set D of formulas is $\mathbb{SWKH}$-inconsistent (or just inconsistent) means that there is a finite subset $D' \subseteq D$ such that $\vdash \neg \bigwedge D'$, where $\bigwedge D' := \bigwedge_{\varphi \in D'} \varphi$ if $D' \neq \emptyset$ and $\bigwedge_{\varphi \in \emptyset} \varphi := \top$. To say that a set of formulas is $\mathbb{SWKH}$-consistent (or just consistent) means that the set of formulas is not inconsistent. Consistency or inconsistency of a formula refers to the consistency or inconsistency of the singleton set containing the formula.*

Axioms	
`TAUT`	Tautologies for propositional logic
`DISTU`	$(\mathcal{U}p \wedge \mathcal{U}(p \to q)) \to \mathcal{U}q$
`TU`	$\mathcal{U}p \to p$
`4WKhU`	$\mathcal{K}h^{\mathtt{w}}(p, q) \to \mathcal{U}\mathcal{K}h^{\mathtt{w}}(p, q)$
`5WKhU`	$\neg\mathcal{K}h^{\mathtt{w}}(p, q) \to \mathcal{U}\neg\mathcal{K}h^{\mathtt{w}}(p, q)$
`EMPWKh`	$\mathcal{U}(p \to q) \to \mathcal{K}h^{\mathtt{w}}(p, q)$
`UWKh`	$(\mathcal{U}(p' \to p) \wedge \mathcal{U}(q \to q') \wedge \mathcal{K}h^{\mathtt{w}}(p, q)) \to \mathcal{K}h^{\mathtt{w}}(p', q')$
Rules	
`MP`	$\dfrac{\varphi, \varphi \to \psi}{\psi}$
`NECU`	$\dfrac{\varphi}{\mathcal{U}\varphi}$
`SUB`	$\dfrac{\varphi}{\varphi[\psi/p]}$

Table 4.1: System $\mathbb{SWKH}$

All the axioms here except `UWKh` are also axioms in the axiomatic system addressed in Wang (2015a), where `UWKh` is deducible from the composition axiom. As observed in Example 4.2.5, the composition axiom is not valid by our semantics. This means that the system here is strictly weaker than Wang (2015a)'s system, which is in line with the fact that here knowledge-how is interpreted in a weaker way. However, even though the system is weaker, the proof of its completeness is non-trivial. We will explain why in the next section.

Proposition 4.2.7 $\vdash \mathcal{U}\chi \wedge \mathcal{U}\psi \to \mathcal{U}(\chi \wedge \psi)$

PROOF
(1) $\vdash \chi \to (\psi \to (\chi \wedge \psi))$ by propositional logic
(2) $\vdash \mathcal{U}(\chi \to (\psi \to (\chi \wedge \psi)))$ by Rule `NECU`
(3) $\vdash \mathcal{U}\chi \to \mathcal{U}(\psi \to (\chi \wedge \psi))$ by Axiom `DISTU`
(4) $\vdash \mathcal{U}(\psi \to (\chi \wedge \psi)) \to (\mathcal{U}\psi \to \mathcal{U}(\chi \wedge \psi))$ by Axiom `DISTU`
(5) $\vdash \mathcal{U}\chi \to (\mathcal{U}\psi \to \mathcal{U}(\chi \wedge \psi))$ by (3) and (4)

(6) $\vdash \mathcal{U}\chi \wedge \mathcal{U}\psi \to \mathcal{U}(\chi \wedge \psi)$ by propositional logic □

Next, we show that $\mathbb{SWKH}$ is sound with respect to the standard semantics.

Proposition 4.2.8 $\vDash \mathcal{U}(p' \to p) \wedge \mathcal{U}(q \to q') \wedge \mathcal{K}h^{\mathtt{w}}(p, q) \to \mathcal{K}h^{\mathtt{w}}(p', q')$

PROOF Assuming that $\mathcal{M}, s \vDash \mathcal{U}(p' \to p) \wedge \mathcal{U}(q \to q') \wedge \mathcal{K}h^{\mathtt{w}}(p, q)$, we need to show that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{w}}(p', q')$. Since $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{w}}(p, q)$, it follows that there exists $\sigma \in Act^*$ such that for each $w \in [\![p]\!]$ and each $t \in \mathtt{TermiSs}(w, \sigma)$ we have $\mathcal{M}, t \vDash q$ $(*)$. In order to show $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{w}}(p', q')$, we only need to show that $\mathcal{M}, t' \vDash q'$ for each $w' \in [\![p']\!]$ and each $t' \in \mathtt{TermiSs}(w', \sigma)$.

Given $w' \in [\![p']\!]$, it follows by $\mathcal{M}, s \vDash \mathcal{U}(p' \to p)$ that $w' \in [\![p]\!]$. Due to $(*)$, we have that for each $t' \in \mathtt{TermiSs}(w', \sigma)$: $\mathcal{M}, t' \vDash q$, namely $t' \in [\![q]\!]$. Moreover, since $\mathcal{M}, s \vDash \mathcal{U}(q \to q')$, we have $[\![q]\!] \subseteq [\![q']\!]$. Therefore, we have that $t' \in [\![q']\!]$, namely $\mathcal{M}, t' \vDash q'$, for each $t' \in \mathtt{TermiSs}(w', \sigma)$. Thus, we have that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathtt{w}}(p', q')$. □

Since $\mathcal{U}$ is a universal modality, `DISTU`, `TU` and `EMPWKh` are obviously valid. Because the modality $\mathcal{K}h^{\mathtt{w}}$ is not local, it is easy to show that `4WKhU` and `5WKhU` are valid. Along with Proposition 4.2.8, we have that all axioms are valid. Moreover, due to a standard argument in modal logic, we know that the rules `MP`, `NECU` and `SUB` preserve a formula's validity. Therefore, the soundness of $\mathbb{SWKH}$ follows immediately.

Theorem 4.2.9 (Soundness) $\mathbb{SWKH}$ *is sound on the standard semantics.*

4.3 Deductive completeness

This section will show that $\mathbb{SWKH}$ is complete with respect to the standard semantics. For the same reason as in Wang (2015a), we will build a canonical model for a given maximally consistent set (MCS). The reason is that the semantics of $\mathcal{K}h^{\mathtt{w}}$ formulas does not depend on the current state. Thus if they are true, they are true everywhere in the model. It follows that we cannot build a *single* canonical model to realize all the consistent sets of $\mathcal{L}_{\text{KHW}}$ formulas simultaneously because both $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi)$ and $\neg\mathcal{K}h^{\mathtt{w}}(\psi, \varphi)$ can be consistent. Instead, for each maximally consistent set of $\mathcal{L}_{\text{KHW}}$ formulas, we build a separate canonical model.

However, the canonical model in Wang (2015a) does not work here because the composition axiom is not valid here, just as shown in Example 4.2.5. We need a new method to construct the canonical model. The canonical model here is also based on a given maximally consistent set Γ, but there are some critical differences. First, the state of the canonical model is a pair consisting of a maximally consistent set and a marker. The marker plays an important role in defining the binary relations of actions. Second, each knowing-how formula is realized by a weak conformant plan consisting of two actions.

Definition 4.3.1 *We say that a set Δ of formulas is* maximally consistent *in $\mathcal{L}_{KHW}$ if Δ is consistent, and any set of formulas properly containing Γ is inconsistent. If Δ is a maximally consistent set of formulas then we say it is an MCS.*

Let Γ be an MCS in $\mathcal{L}_{\text{KHW}}$. In the following, we will build a canonical model for Γ. We first prepare ourselves with some auxiliary notions and some handy propositions.

Given a set of $\mathcal{L}_{\text{KHW}}$ formulas Δ, let $\Delta|_{\mathcal{K}h^{\mathtt{w}}}$ and $\Delta|_{\neg\mathcal{K}h^{\mathtt{w}}}$ be the collections of its positive and negative $\mathcal{K}h^{\mathtt{w}}$ formulas:

$$\Delta|_{\mathcal{K}h^{\mathtt{w}}} = \{\theta \in \Delta \mid \theta \text{ is of the form } \mathcal{K}h^{\mathtt{w}}(\psi, \varphi)\};$$
$$\Delta|_{\neg\mathcal{K}h^{\mathtt{w}}} = \{\theta \in \Delta \mid \theta \text{ is of the form } \neg\mathcal{K}h^{\mathtt{w}}(\psi, \varphi)\}.$$

Definition 4.3.2 *Let Φ_Γ be the set of all MCS Δ such that $\Delta|_{\mathcal{K}h^{\mathtt{w}}} = \Gamma|_{\mathcal{K}h^{\mathtt{w}}}$.*

Note that Φ_Γ is the set of all MCSs that share the same $\mathcal{K}h^{\mathtt{m}}$ formulas with Γ. The canonical model for Γ will be based on the MCSs in Φ_Γ. Since every $\Delta \in \Phi_\Gamma$ is maximally consistent, the following proposition shows an obvious property of Φ_Γ.

Proposition 4.3.3 *For each $\Delta \in \Phi_\Gamma$, we have $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi) \in \Gamma$ if and only if $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi) \in \Delta$ for all $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi) \in \mathcal{L}_{KHW}$.*

By a standard argument of Lindenbaum's lemma (cf. Blackburn et al. (2001)), we have the following proposition.

Proposition 4.3.4 *If Δ is consistent then there is an MCS Γ such that $\Delta \subseteq \Gamma$.*

The following proposition reveals a crucial property of Φ_Γ, which will be used repeatedly later on.

Proposition 4.3.5 *If $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$ then $\mathcal{U}\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$.*

PROOF Suppose $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$, then by the definition of Φ_Γ, $\neg\varphi$ is not consistent with $\Gamma|_{\mathcal{K}h^{\mathtt{w}}} \cup \Gamma|_{\neg\mathcal{K}h^{\mathtt{w}}}$, for otherwise $\Gamma|_{\mathcal{K}h^{\mathtt{w}}} \cup \Gamma|_{\neg\mathcal{K}h^{\mathtt{w}}} \cup \{\neg\varphi\}$ can be extended into a maximally consistent set in Φ_Γ due to Proposition 4.3.4, which contradicts the assumption that $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. Thus there are $\mathcal{K}h^{\mathtt{w}}(\psi_1, \varphi_1), \dots, \mathcal{K}h^{\mathtt{w}}(\psi_k, \varphi_k) \in \Gamma|_{\mathcal{K}h^{\mathtt{w}}}$ and $\neg\mathcal{K}h^{\mathtt{w}}(\psi'_1, \varphi'_1), \dots, \neg\mathcal{K}h^{\mathtt{w}}(\psi'_l, \varphi'_l) \in \Gamma|_{\neg\mathcal{K}h^{\mathtt{w}}}$ such that

$$\vdash \bigwedge_{1 \le i \le k} \mathcal{K}h^{\mathtt{w}}(\psi_i, \varphi_i) \wedge \bigwedge_{1 \le j \le l} \neg\mathcal{K}h^{\mathtt{w}}(\psi'_j, \varphi'_j) \to \varphi.$$

By `NECU`,

$$\vdash \mathcal{U}(\bigwedge_{1 \le i \le k} \mathcal{K}h^{\mathtt{w}}(\psi_i, \varphi_i) \wedge \bigwedge_{1 \le j \le l} \neg\mathcal{K}h^{\mathtt{w}}(\psi'_j, \varphi'_j) \to \varphi).$$

By `DISTU` we have:

$$\vdash \mathcal{U}(\bigwedge_{1 \le i \le k} \mathcal{K}h^{\mathtt{w}}(\psi_i, \varphi_i) \wedge \bigwedge_{1 \le j \le l} \neg\mathcal{K}h^{\mathtt{w}}(\psi'_j, \varphi'_j)) \to \mathcal{U}\varphi.$$

Since $\mathcal{K}h^{\mathtt{w}}(\psi_1, \varphi_1), \dots, \mathcal{K}h^{\mathtt{w}}(\psi_k, \varphi_k) \in \Gamma|_{\mathcal{K}h^{\mathtt{w}}}$, it follows that $\mathcal{U}\mathcal{K}h^{\mathtt{w}}(\psi_1, \varphi_1), \dots, \mathcal{U}\mathcal{K}h^{\mathtt{w}}(\psi_k, \varphi_k) \in \Gamma$ due to `4WKhU` and the fact that Γ is a maximally consistent set. Similarly, we have $\mathcal{U}\neg\mathcal{K}h^{\mathtt{w}}(\psi'_1, \varphi'_1), \dots, \mathcal{U}\neg\mathcal{K}h^{\mathtt{w}}(\psi'_j, \varphi'_j) \in \Gamma$ due to `5WKhU`. By Proposition 4.2.7, it follows that

$$\mathcal{U}(\bigwedge_{1 \le i \le k} \mathcal{K}h^{\mathtt{w}}(\psi_i, \varphi_i) \wedge \bigwedge_{1 \le j \le l} \neg\mathcal{K}h^{\mathtt{w}}(\psi'_j, \varphi'_j)) \in \Gamma.$$

Now it is immediate that $\mathcal{U}\varphi \in \Gamma$. Due to Proposition 4.3.3, $\mathcal{U}\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. $\Box$

Proposition 4.3.6 *Given $\mathcal{K}h^{\mathtt{w}}(\psi,\varphi) \in \Gamma$ and $\Delta \in \Phi_\Gamma$, if $\psi \in \Delta$ then there exists $\Delta' \in \Phi_\Gamma$ such that $\varphi \in \Delta'$.*

PROOF Assuming $\mathcal{K}h^{\mathtt{w}}(\psi,\varphi) \in \Gamma$ and $\psi \in \Delta \in \Phi_\Gamma$, if there does not exist $\Delta' \in \Phi_\Gamma$ such that $\varphi \in \Delta'$, it means that $\neg\varphi \in \Delta'$ for all $\Delta' \in \Phi_\Gamma$. It follows by Proposition 4.3.5 that $\mathcal{U}\neg\varphi \in \Gamma$, and then $\mathcal{U}(\varphi \to \bot) \in \Gamma$. Since $\mathcal{U}(\varphi \to \bot)$ and $\mathcal{K}h^{\mathtt{w}}(\psi,\varphi) \in \Gamma$, it follows by `UWKh` that $\mathcal{K}h^{\mathtt{w}}(\psi,\bot) \in \Gamma$, namely $\mathcal{U}\neg\psi \in \Gamma$. By Proposition 4.3.3, we have that $\mathcal{U}\neg\psi \in \Delta$. It follows by `TU` that $\neg\psi \in \Delta$. This is in contradiction with $\psi \in \Delta$. Therefore, there exists $\Delta' \in \Phi_\Gamma$ such that $\varphi \in \Delta'$. □

Next, we will construct the canonical model for the MCS Γ. It is crucial first to understand the ideas behind the canonical model construction. Besides satisfying $\mathcal{K}h^{\mathtt{w}}(\psi,\varphi)$, the canonical model also needs to meet the following two requirements.

(1) Generally, $\mathcal{K}h^{\mathtt{w}}(\psi,\varphi)$ cannot be satisfied by a one-step plan. Otherwise, the canonical model will always satisfy the formula that $\mathcal{K}h^{\mathtt{w}}(p,\neg p) \wedge \mathcal{K}h^{\mathtt{w}}(\neg p,q) \to \mathcal{K}h^{\mathtt{w}}(p,q)$ which is not a valid formula. Therefore, in the canonical model, $\mathcal{K}h^{\mathtt{w}}(\psi,\varphi)$ will be realized by a two-step plan $\langle\psi,\psi\varphi\rangle\langle\psi\varphi,\varphi\rangle$. If we already reach a φ-state by the first step $\langle\psi,\psi\varphi\rangle$, we do not need to go further anymore. If we arrive at a $\neg\varphi$-state by $\langle\psi,\psi\varphi\rangle$, then we need to make sure that doing the second step $\langle\psi\varphi,\varphi\rangle$ on this state will achieve only φ-states.

(2) If $\langle\psi,\psi\varphi\rangle\langle\psi\varphi,\varphi\rangle$ is a weak conformant plan for $\mathcal{K}h^{\mathtt{m}}(\psi,\varphi)$, then $\langle\psi,\psi\varphi\rangle$ must be executable on at least one $\neg\psi$-state. The reason is that if $\langle\psi,\psi\varphi\rangle$ is only executable at ψ-states then the canonical model will always satisfy $\mathcal{K}h^{\mathtt{w}}(\psi,\varphi) \to \mathcal{K}h^{\mathtt{w}}(\psi \vee \varphi,\varphi)$ which is not a valid formula. If we allow $\langle\psi,\psi\varphi\rangle$ also executable at $\neg\psi$-states, we must treat the step from ψ-states and $\neg\psi$-states differently. Otherwise, the canonical model will always satisfy $\mathcal{K}h^{\mathtt{w}}(\psi,\varphi) \to \mathcal{K}h^{\mathtt{w}}(\top,\varphi)$. Our method is that the step $\langle\psi,\psi\varphi\rangle$ starting from ψ-states will reach only states marked with $\psi\varphi$. This is why we include $\psi\varphi$ markers in the building blocks of the canonical model besides maximally consistent set.[3]

Definition 4.3.7 (Canonical Model) *The canonical model for Γ is a quadruple $\mathcal{M}^c_\Gamma = \langle S^c, Act_\Gamma, R^c, V^c\rangle$ where:*

- $S^c = \{(\Delta,\psi\varphi) \mid \Delta \in \Phi_\Gamma, \mathcal{K}h^{\mathtt{w}}(\psi,\varphi) \in \Gamma\}$. *We write the pair in S as $w, v, \cdots$, and refer to the first entry of $w \in S$ as $\mathtt{L}(w)$, to the second entry as $\mathtt{R}(w)$;*
- $Act_\Gamma = \{\langle\psi,\psi\varphi\rangle, \langle\psi\varphi,\varphi\rangle \mid \mathcal{K}h^{\mathtt{w}}(\psi,\varphi) \in \Gamma\}$*;*
- $w \xrightarrow{\langle\psi,\psi\varphi\rangle}_c w' \iff$ *If $\psi \in \mathtt{L}(w)$ then $\mathtt{R}(w') = \psi\varphi$;*
- $w \xrightarrow{\langle\psi\varphi,\varphi\rangle}_c w' \iff \mathtt{R}(w) = \mathtt{R}(w') = \psi\varphi$, $\neg\varphi \in \mathtt{L}(w)$ *and* $\varphi \in \mathtt{L}(w')$*;*
- $p \in V^c(w) \iff p \in \mathtt{L}(w)$.

For each $w \in S$, we also call w a ψ-state if $\psi \in \mathtt{L}(w)$.

[3] In Wang (2015a), the canonical models are much simpler: we just need MCSs and the canonical relations are simply labelled by $\langle\psi,\varphi\rangle$ for $\mathcal{K}h(\psi,\varphi) \in \Gamma$.

Please note that S^c is non-empty because $(\Gamma, \top\top) \in S^c$. We first show that each $\Delta \in \Phi_\Gamma$ appears as $\mathtt{L}(w)$ for some $w \in S^c$.

Proposition 4.3.8 *For each $\Delta \in \Phi_\Gamma$, there exists $w \in S^c$ such that $\mathtt{L}(w) = \Delta$.*

PROOF Since $\vdash \top \to \top$, it follows by NECU that $\vdash \mathcal{U}(\top \to \top)$. Thus, we have $\mathcal{U}(\top \to \top) \in \Gamma$. It follows by EMPWKh that $\mathcal{K}h^{\mathtt{W}}(\top, \top) \in \Gamma$. Thus, we have that $(\Delta, \top\top) \in S^c$. □

Since $\Gamma \in \Phi_\Gamma$, it follows by Proposition 4.3.8 that $S^c \neq \emptyset$.

The following proposition indicates that in the canonical model each knowing-how formula $\mathcal{K}h(\psi, \varphi)$ is realized by a plan whose length is no more than 2.

Proposition 4.3.9 *If there is $\sigma = a_1 \cdots a_n \in Act_\Gamma^*$ where $n \geq 3$ such that for each ψ-state $w \in S^c$ and each state $t \in \mathtt{TermiSs}(w, \sigma)$ we have $\varphi \in \mathtt{L}(t)$, then there exists $\sigma' = a'_1 \cdots a'_k \in Act_\Gamma^*$ where $k \leq n-1$ such that for each ψ-state $w \in S^c$ and each state $t \in \mathtt{TermiSs}(w, \sigma)$ we have $\varphi \in \mathtt{L}(t)$.*

PROOF Let s be a ψ-state such that $s \xrightarrow{\sigma}_c t$ for some $t \in S^c$. If there is no such a state, it is easy to show that $\mathtt{TermiSs}(w, \sigma) = \mathtt{TermiSs}(w, \sigma_{n-1})$ for each ψ-state w, and then $\sigma' = a_1 \cdots a_{n-1}$ satisfies the conditions. Next we proceed the proof by considering the following two cases according to the form of a_1.

- $a_1 = \langle \psi_1\varphi_1, \varphi_1 \rangle$ We will show that $\sigma' = \epsilon$ satisfies that for each ψ-state $w \in S^c$ and each state $t \in \mathtt{TermiSs}(w, \sigma')$ we have $\varphi \in \mathtt{L}(t)$. We only need to show that $\psi \to \varphi \in \Delta$ for each $\Delta \in \Phi_\Gamma$. If not, there exists $\Delta' \in \Phi_\Gamma$ such that $\{\psi, \neg\varphi\} \subset \Delta'$. Let χ be a formula such that $\vdash \chi \leftrightarrow \psi_1$ and $\chi \neq \psi_1$. Since $\vdash \chi \to \top$, it follows by NECU and EMPWKh that $\mathcal{K}h^{\mathtt{W}}(\chi, \top) \in \Gamma$. Then we have a ψ-state $w' = (\Delta', \chi\top) \in S^c$. Since $\chi \neq \psi_1$, a_1 is not executable on w', and then we have $\{w'\} = \mathtt{TermiSs}(w', \sigma)$. Since $\neg\varphi \in \mathtt{L}(w')$, this is in contradiction with our assumption. Thus we have $\psi \to \varphi \in \Delta$ for each $\Delta \in \Phi_\Gamma$.

- $a_1 = \langle \psi_1, \psi_1\varphi_1 \rangle$ There are two cases based on the form of a_2:

  - $a_2 = \langle \psi_2, \psi_2\varphi_2 \rangle$ There are two cases: $\mathcal{U}\neg\psi_2 \in \Gamma$ or not.

    * There is no $\Delta \in \Phi_\Gamma$ such that $\neg\psi_2 \in \Delta$. Let $\sigma' = a_2 \cdots a_n$. Given a ψ-state w, next we will show that if $t \in \mathtt{TermiSs}(w, \sigma')$ then $\varphi \in \mathtt{L}(t)$. For each $t \in \mathtt{TermiSs}(w, \sigma')$ we have $w \xrightarrow{a_2}_c t' \cdots \xrightarrow{a_k}_c t$ where $2 \leq k \leq n$. Duo to $\psi_2 \in \mathtt{L}(w)$, it follows by the definition of $\xrightarrow{a_2}_c$ that $t' = (\Delta', \psi_2\varphi_2)$ for some $\Delta' \in \Phi_\Gamma$. What is more, for each s' with $s \xrightarrow{a_1}_c s'$, we have $s' \xrightarrow{a_2}_c t'$ duo to $\psi_2 \in \mathtt{L}(s')$. (Please note s is the state we mentioned at the beginning of the proof.) It follows that $t \in \mathtt{TermiSs}(s, \sigma)$. It follows by assumption that $\varphi \in \mathtt{L}(t)$.
    * There exists $\Delta \in \Phi_\Gamma$ such that $\neg\psi_2 \in \Delta$. Let $s' = (\Delta, \psi_1\varphi_1)$ for some $\neg\psi_2 \in \Delta$. It follows that $s \xrightarrow{a_1}_c s'$ and $s' \xrightarrow{a_2}_c t'$ for each $t' \in S^c$. Let $\sigma' = a_3 \cdots a_n$. Given a ψ-state w, we have $s \xrightarrow{a_1}_c s' \xrightarrow{a_2}_c w$. For each $t \in \mathtt{TermiSs}(w, \sigma')$, it follows that $t \in \mathtt{TermiSs}(s, \sigma)$. Thus, we have $\varphi \in \mathtt{L}(t)$. Therefore, σ' satisfies the conditions.

- $a_2 = \langle \psi_2\varphi_2, \varphi_2 \rangle$ There are two cases: $\mathcal{U}(\psi \to \psi_1) \in \Gamma$ or not.
  - There exists $\Delta \in \Phi_\Gamma$ such that $\psi, \neg\psi_1 \in \Delta$. In this case, it must be that $\varphi \in \Delta'$ for each $\Delta' \in \Phi_\Gamma$. If not, let $t = (\Delta', \psi_2'\varphi_2)$ where $\neg\varphi \in \Delta'$, $\vdash \psi_2 \leftrightarrow \psi_2'$ and $\psi_2 \neq \psi_2$. Let w be a state such that $\{\psi, \neg\psi_1\} \subset \mathtt{L}(w)$. It follows that $w \xrightarrow{a_1}_c t$ and a_2 is not executable at t. Thus, $t \in \mathtt{TermiSs}(w, \sigma)$. By assumption, we have $\varphi \in \mathtt{L}(t)$. Contradiction. Therefore, we have $\varphi \in \Delta'$ for each $\Delta' \in \Phi_\Gamma$. Then, $\sigma' = \epsilon$ satisfies that for each ψ-state w and each $t \in \mathtt{TermiSs}(w, \sigma')$ we have $\varphi \in \mathtt{L}(t)$.
  - There is no $\Delta \in \Phi_\Gamma$ such that $\psi, \neg\psi_1 \in \Delta$. Thus, we have $\psi_1 \in \mathtt{L}(s)$. (Please note s is the state mentioned at the beginning of the proof.) Since $s \xrightarrow{a_1}_c s_1 \xrightarrow{a_2}_c s_2$ for some $s_1, s_2 \in S^c$, it follows by the definition of a_1 and a_2 that $\psi_1 = \psi_2$ and $\varphi_1 = \varphi_2$. Firstly we will show that $\varphi_1 \to \varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. If not, there is $\Delta' \in \Phi_\Gamma$ such that $\{\varphi_1, \neg\varphi\} \subset \Delta'$. Let $t = (\Delta', \psi_1\varphi_1)$ then we have $s \xrightarrow{a_1}_c t$ and a_2 is not executable at t. Thus, we have $t \in \mathtt{TermiSs}(s, \sigma)$. Since $\neg\varphi \in \mathtt{L}(t)$, it is in contradiction with our assumption. Therefore, we have $\varphi_1 \to \varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$ $(\star\star)$.
    Let $\sigma' = a_1a_2$. Given a ψ-state w, for each $t \in \mathtt{TermiSs}(w, \sigma')$, there are two cases: $w \xrightarrow{a_1}_c t$ and a_2 is not executable at t, or $w \xrightarrow{a_1}_c w' \xrightarrow{a_2}_c t$. Both of them have that $\varphi_1 \in \mathtt{L}(t)$. It follows by $(\star\star)$ that $\varphi \in \mathtt{L}(t)$.

□

Now we are ready to prove the truth lemma.

Lemma 4.3.10 (Truth Lemma) *For each φ, we have $\mathcal{M}^{c}_{\Gamma}, w \vDash \varphi$ iff $\varphi \in \mathtt{L}(w)$.*

PROOF Boolean cases are trivial, and we only focus on the case of $\mathcal{K}h^{\mathtt{W}}(\psi, \varphi)$.

Left to Right: Supposing $\mathcal{K}h^{\mathtt{W}}(\psi, \varphi) \in \mathtt{L}(w)$, let $a_1 = \langle \psi, \psi\varphi \rangle$ and $a_2 = \langle \psi\varphi, \varphi \rangle$, then we will show that $\mathcal{M}^{c}_{\Gamma}, t \vDash \varphi$ for each $w \in [\![\psi]\!]$ and each $t \in \mathtt{TermiSs}(w, a_1a_2)$. Given $w \in [\![\psi]\!]$ and $t \in \mathtt{TermiSs}(w, a_1a_2)$, we will show that $\mathcal{M}^{c}_{\Gamma}, t \vDash \varphi$. By IH, we only need to show that $\varphi \in \mathtt{L}(t)$. For $t \in \mathtt{TermiSs}(w, a_1a_2)$, there are two cases:

- $w \xrightarrow{a_1}_c t$ and a_2 is not executable at t. Since $w \in [\![\psi]\!]$, it follows by IH that $\psi \in \mathtt{L}(w)$.By the definition of $\xrightarrow{a_1}_c$, we have $\mathtt{R}(t) = \psi\varphi$. Due to $\psi \in \mathtt{L}(w)$ and $\mathcal{K}h^{\mathtt{W}}(\psi, \varphi) \in \Gamma$, it follows by Proposition 4.3.6 that $\varphi \in \Delta'$ for some $\Delta' \in \Phi_\Gamma$. Thus, $(\Delta', \psi\varphi)$ is a state in S^c. Then, there is only one reason left for a_2 not executable at t, that is, $\neg\varphi \notin \mathtt{L}(t)$. Therefore, we have $\varphi \in \mathtt{L}(t)$.

- $w \xrightarrow{a_1}_c w' \xrightarrow{a_2}_c t$ for some $w' \in S^c$. By the definition of $\xrightarrow{a_2}_c$, it follows that $\varphi \in \mathtt{L}(t)$.

Right to Left: If $\mathcal{M}^{c}_{\Gamma}, w \vDash \varphi$, we assume that $\sigma = a_1 \cdots a_n \in Act^*_\Gamma$ is the shortest action sequence such that $\mathcal{M}^{c}_{\Gamma}, t \vDash \varphi$ for each $w \in [\![\psi]\!]$ and each $t \in \mathtt{TermiSs}(w, \sigma)$. It follows by Proposition 4.3.9 and IH that $n \leq 2$. Let us consider the following two cases: $n = 0$ or $n > 0$.

- $n = 0$ It means $\sigma = \epsilon$. It follows by IH that $\psi \in \Delta$ implies $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. Therefore, we have $\psi \to \varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. It follows by Proposition 4.3.5 that $\mathcal{U}(\psi \to \varphi) \in \Gamma$. By $\mathtt{EMPWKh}$, we have that $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi) \in \Gamma$. It follows by Proposition 4.3.3 that $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi) \in \mathtt{L}(w)$.

- $n > 0$ There are three cases.

  – a_1 cannot be of the form $\langle \psi_1\varphi_1, \varphi_1\rangle$. If $a_1 = \langle \psi_1\varphi_1, \varphi_1\rangle$, we can show that $\psi \to \varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. If not, there is $\Delta \in \Phi_\Gamma$ such that $\{\psi, \neg\varphi\} \subset \Delta$. Let χ be a formula such that $\vdash \psi_1 \leftrightarrow \chi$ and $\psi_1 \neq \chi$. Due to $\mathcal{U}(\chi \to \top) \in \Gamma$, it follows that $\mathcal{K}h^{\mathtt{w}}(\chi, \top) \in \Gamma$. Thus $t = (\Delta, \chi\top)$ is a state in S^c. By IH, we have $\mathcal{M}^c_\Gamma, t \vDash \psi$. Since a_1 is not executable at t, we have $t \in \mathtt{TermiSs}(t, \sigma)$. Therefore, $\mathcal{M}^c_\Gamma, t \vDash \varphi$. By IH, we have $\varphi \in \mathtt{L}(t)$. Contradiction. Thus, we have $\psi \to \varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. By IH, we have that $\mathcal{M}^c_\Gamma, t \vDash \varphi$ for each $w' \in [\![\psi]\!]$ and each $t \in \mathtt{TermiSs}(w', \epsilon)$. This is in contradiction with our assumption that σ is the shortest and $n > 0$.

  – a_n cannot be of the form $\langle \psi_n, \psi_n\varphi_n\rangle$. If $a_n = \langle \psi_n, \psi_n\varphi_n\rangle$, we can show that $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. Otherwise, there is $\Delta' \in \Phi_\Gamma$ such that $\neg\varphi \in \Delta'$. let $t = (\Delta', \psi_n\varphi_n)$. Since σ is the shortest, it follows that there is a state w' such that $\mathcal{M}^c_\Gamma, w' \vDash \psi$, $w' \xrightarrow{\sigma_{n-1}}_c v$ and a_n is executable at v. Then we have $v \xrightarrow{a_n}_c t$. Thus, we have that $w' \xrightarrow{\sigma}_c t$. It follows that $t \in \mathtt{TermiSs}(w', \sigma)$. Thus, we have $\mathcal{M}^c_\Gamma, t \vDash \varphi$. It follows by IH that $\varphi \in \mathtt{L}(t)$. Contradiction. Thus, we have $\varphi \in \Delta$ for all $\Delta \in \Phi_\Gamma$. By IH, we have that $\mathcal{M}^c_\Gamma, t \vDash \varphi$ for each $w' \in [\![\psi]\!]$ and each $t \in \mathtt{TermiSs}(w', \epsilon)$. This is in contradiction with our assumption that σ is the shortest and $n > 0$.

  – $a_1 = \langle \psi_1, \psi_1\varphi_1\rangle$ and $a_2 = \langle \psi_2\varphi_2, \varphi_2\rangle$.
  Firstly, we show that there is no $\Delta \in \Phi_\Gamma$ such that $\{\psi, \neg\psi_1\} \subset \Delta$, namely $\mathcal{U}(\psi \to \psi_1) \in \Gamma$. If there is such a MCS Δ, let w be state such that $\mathtt{L}(w) = \Delta$. It follows that $w \xrightarrow{a_1}_c t$ for all $t \in S^c$. Then, it must be that $\varphi \in \Delta'$ for all $\Delta' \in \Phi_\Gamma$. Otherwise, let $t = (\Delta', \psi'_2\varphi_2)$ where $\neg\varphi \in \Delta'$, $\vdash \psi'_2 \leftrightarrow \psi_2$ and $\psi'_2 \neq \psi_2$. It follows that $w \xrightarrow{a_1}_c t$ and a_2 is not executable at t, namely $t \in \mathtt{TermiSs}(w, a_1a_2)$. It follows that $\mathcal{M}^c_\Gamma, t \vDash \varphi$, and then by IH that $\varphi \in \mathtt{L}(t) = \Delta'$. Contradiction. Thus, we have $\varphi \in \Delta'$ for all $\Delta' \in \Phi_\Gamma$. Then, $\sigma' = \epsilon$ satisfies that for each $w \in [\![\psi]\!]$ and each $t \in \mathtt{TermiSs}(w, \epsilon)$ we have $\mathcal{M}^c_\Gamma, t \vDash \varphi$. This is in contradiction with $\sigma = a_1a_2$ is the shortest one. Therefore, we have there is no $\Delta \in \Phi_\Gamma$ such that $\{\psi, \neg\psi_1\} \subset \Delta$. It, then, follows by Proposition 4.3.5 that $\mathcal{U}(\psi \to \psi_1) \in \Gamma$.

  Next, we will show that there is no $\Delta \in \Phi_\Gamma$ such that $\{\varphi_1, \neg\varphi\} \subset \Delta$, namely $\mathcal{U}(\varphi_1 \to \varphi) \in \Gamma$. Since σ is the shortest, we assume that w', t_1, t_2 are states such that $w' \xrightarrow{a_1}_c t_1 \xrightarrow{a_2}_c t_2$ and $w' \in [\![\psi]\!]$. It follows that $\psi_1 = \psi_2$ and $\varphi_1 = \varphi_2$. If there is $\Delta \in \Phi_\Gamma$ such that $\{\varphi_1, \neg\varphi\} \subset \Delta$, let $t = (\Delta, \psi_1\varphi_1)$. It follows that $w' \xrightarrow{a_1}_c t_1 \xrightarrow{a_2}_c t$. Thus, we have $t \in \mathtt{TermiSs}(w', a_1a_2)$. It follows that $\mathcal{M}^c_\Gamma, t \vDash \varphi$, and then by IH that $\varphi \in \mathtt{L}(t) = \Delta$. Contradiction. Thus, we have shown that there is no $\Delta \in \Phi_\Gamma$ such that $\{\varphi_1, \neg\varphi\} \subset \Delta$, and then by Proposition 4.3.5 that $\mathcal{U}(\varphi_1 \to \varphi) \in \Gamma$. Due to $a_1 \in Act_\Gamma$, we have $\mathcal{K}h^{\mathtt{w}}(\psi_1, \varphi_1) \in \Gamma$. Since we have shown that $\mathcal{U}(\psi \to \psi_1) \in \Gamma$ and $\mathcal{U}(\varphi_1 \to \varphi) \in \Gamma$, it follows by $\texttt{UWKh}$ that $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi) \in \Gamma$, and then $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi) \in \mathtt{L}(w)$.

□

The key of the completeness proof is to show that each $\mathbb{SWKH}$-consistent set is satisfiable. Due to a standard Lindenbaum-like argument, each $\mathbb{SWKH}$-consistent set of formulas can be extended to a maximally consistent set Γ. Due to the truth lemma, we have that $\mathcal{M}^c_\Gamma, (\Gamma, \top\top) \vDash \Gamma$. The completeness of $\mathbb{SWKH}$ follows immediately.

Theorem 4.3.11 (Strong Completeness) $\mathbb{SWKH}$ *is strongly complete on the standard semantics.*

4.4 Decidability

This section will show that the problem whether a formula φ is valid with respect to the standard semantics is decidable. The strategy is that we firstly define a nonstandard semantics and then show that φ is valid with respect to the standard semantics if and only if φ is valid with respect to the nonstandard semantics. Next, we show that φ has a bounded small model if φ is satisfiable with respect to the nonstandard model.

Definition 4.4.1 (Nonstandard semantics) *Given a pointed model $\mathcal{M}, s$ and a formula φ, we write $\mathcal{M}, s \Vdash \varphi$ to mean that φ is true at $\mathcal{M}, s$ with respect to the nonstandard semantics $\Vdash$. The nonstandard semantics $\Vdash$ is defined by the following induction on formula construction.*

$\mathcal{M}, s \Vdash \top$		*always*
$\mathcal{M}, s \Vdash p$	$\iff$	$s \in V(p)$.
$\mathcal{M}, s \Vdash \neg\varphi$	$\iff$	$\mathcal{M}, s \nVdash \varphi$.
$\mathcal{M}, s \Vdash \varphi \wedge \psi$	$\iff$	$\mathcal{M}, s \Vdash \varphi$ *and* $\mathcal{M}, s \Vdash \psi$.
$\mathcal{M}, s \Vdash \mathcal{K}h^{\mathtt{w}}(\psi, \varphi)$	$\iff$	*there exists* $a \in Act^\bullet$ *such that for all* $\mathcal{M}, u \Vdash \psi$: *a is executable at u and* $\mathcal{M}, v \Vdash \varphi$ *for all* $v \in R_a(u)$

where $Act^\bullet = Act \cup \{\epsilon\}$. To say φ is valid with respect to the nonstandard semantics, written $\Vdash \varphi$, means $\mathcal{M}, s \Vdash \varphi$ for all pointed model $\mathcal{M}, s$.

In this nonstandard semantics, the knowledge-how is interpreted almost the same as in Moore's first interpretation (I). The only difference is that the witness action for the knowledge-how might be epsilon ϵ. Intuitively, this means that if φ is true in each ψ-state then we know how to achieve φ given ψ trivially by doing nothing.

Let $\mathcal{M}, s \Vdash \mathcal{U}\varphi$ be defined as $\mathcal{M}, u \Vdash \varphi$ for all $u \in S$. It is easy to show that

$$\mathcal{M}, s \Vdash \mathcal{U}\varphi \iff \mathcal{M}, s \Vdash \mathcal{K}h^{\mathtt{w}}(\neg\varphi, \bot)$$

To show that $\vDash \varphi$ if and only if $\Vdash \varphi$, since the axiom system $\mathbb{SWKH}$ is sound and complete with respect to the standard semantics, we only need to show that $\mathbb{SWKH}$ is also sound and complete on the nonstandard semantics.

Since $\mathcal{K}h^{\mathtt{w}}$ is also a universal modality, it is easy to verify that $\mathbb{SWKH}$ is sound on the nonstandard semantics.

Proposition 4.4.2 (Soundness w.r.t. nonstandard semantics) *If* $\vdash \varphi$ *then* $\Vdash \varphi$.

Next, we will show that $\mathbb{SWKH}$ is complete on the nonstandard semantics. The key is to construct a model for a given consistent formula φ such that φ is satisfiable in the model on the nonstandard semantics.

Here are some notions before we construct the model for φ. We use $Sub(\varphi)$ to denote the set of all sub-formulas of φ. Let $\sim\psi = \chi$ if $\psi = \neg\chi$, otherwise, $\sim\psi = \neg\psi$. It is obvious that $\vdash \neg\psi \leftrightarrow \sim\psi$. We use $Sub^+(\varphi)$ to denote the set $Sub(\varphi) \cup \{\sim\psi \mid \psi \in Sub(\varphi)\}$.

Similar to the completeness of $\mathbb{SWKH}$ to the standard semantics, the model we are going to construct here is based on maximally consistent sets in the set $Sub^+(\varphi)$.

Proposition 4.4.3 *If* Γ *is a consistent subset of* $Sub^+(\varphi)$ *then there exists an MCS* B *in* $Sub^+(\varphi)$ *such that* $\Gamma \subseteq B$.

PROOF Let Φ be the set of all MCSs in $Sub^+(\varphi)$. It is easy to show that $\vdash \bigvee_{A\in\Phi} A$. Assume that there is no $A \in \Phi$ such that $\Gamma \subseteq A$. It follows that there is $\psi \in \Gamma$ such that $\sim\psi \in A$. Thus, we have $A \vdash \neg\psi$. Since $\vdash \neg\psi \to \neg\bigwedge\Gamma$, it follows that $A \vdash \neg\bigwedge\Gamma$. Therefore, $\vdash \bigvee_{A\in\Phi} A \to \neg\bigwedge\Gamma$, and then we have $\vdash \neg\bigwedge\Gamma$. This is in contradiction with Γ being consistent. Thus, we have that there exists $A \in \Phi$ such that $\Gamma \subseteq A$. $\square$

As we did in Section 4.3, here we will construct the model based on a given MCS in $Sub^+(\varphi)$. Let A be an MCS in $Sub^+(\varphi)$. We use Θ^A to denote $A|_{\mathcal{K}h^{\mathtt{w}}} \cup A|_{\neg\mathcal{K}h^{\mathtt{w}}}$. The model based on A is defined in the following.

Definition 4.4.4 *The model* $\mathcal{M}^A = \langle S^A, Act^A, R^A, V^A\rangle$ *is defined as follows.*

- $S^A = \{B \text{ is an MCS in } Sub^+(\varphi) \mid (B|_{\mathcal{K}h^{\mathtt{w}}} \cup B|_{\neg\mathcal{K}h^{\mathtt{w}}}) = \Theta^A\}$;
- $Act^A = \{\langle\chi,\psi\rangle \mid \mathcal{K}h^{\mathtt{w}}(\chi,\psi) \in \Theta^A\}$;
- $B \xrightarrow{\langle\chi,\psi\rangle} B' \iff \chi \in B$ *and* $\psi \in B'$, *for each* $\langle\chi,\psi\rangle \in Act^A$;
- $p \in V^A(B) \iff p \in B$, *for each* $p \in Sub^+(\varphi)$.

The domain S^A is non-empty because $A \in S^A$. Next, we will show the truth lemma for this model. Before that, we first prepare ourselves with some useful handy propositions.

Proposition 4.4.5 $\Theta^A \vdash \mathcal{U}\bigwedge\Theta^A$

PROOF Let $\bigwedge\Theta^A := \theta_1 \wedge \theta_2$ where $\theta_1 := \mathcal{K}h^{\mathtt{w}}(\chi_1,\psi_1) \wedge \cdots \wedge \mathcal{K}h^{\mathtt{w}}(\chi_n,\psi_n)$ and $\theta_2 := \neg\mathcal{K}h^{\mathtt{w}}(\chi'_1,\psi'_1) \wedge \cdots \wedge \neg\mathcal{K}h^{\mathtt{w}}(\chi'_m,\psi'_m)$. It follows by Axiom `4WKhU` and Proposition 4.2.7 that $\vdash \theta_1 \to \mathcal{U}\theta_1$. It follows by Axiom `5WKhU` and Proposition 4.2.7 that $\vdash \theta_2 \to \mathcal{U}\theta_2$. Again by Proposition 4.2.7, we have $\vdash \theta_1 \wedge \theta_2 \to \mathcal{U}(\theta_1 \wedge \theta_2)$. Thus, we have $\Theta^A \vdash \mathcal{U}\bigwedge\Theta^A$. $\square$

Proposition 4.4.6 *For each* $\psi \in Sub^+(\varphi)$, *if* $\psi \in B$ *for all* $B \in S^A$ *then* $\Theta^A \vdash \mathcal{U}\psi$.

PROOF Since $\psi \in B$ for all $B \in S^A$, it follows that $\Theta^A \cup \{\neg\psi\}$ is inconsistent. It follows that $\vdash \bigwedge \Theta^A \to \psi$. By Rule `NECU`, it follows that $\vdash \mathcal{U}(\bigwedge \Theta^A \to \psi)$. It follows by Axiom `DISTU` that $\vdash \mathcal{U} \bigwedge \Theta^A \to \mathcal{U}\psi$. By Proposition 4.4.5, it follows that Therefore, $\Theta^A \vdash \mathcal{U}\psi$. $\square$

Proposition 4.4.7 *Given $\chi \in Sub^+(\varphi)$ and $B \in S^A$, if $\chi \in B$ implies that $\langle \chi', \psi' \rangle \in Act^A$ is executable at B then we have $\Theta^A \vdash \mathcal{U}(\chi \to \chi')$.*

PROOF Assume that $\Theta^A \cup \{\chi, {\sim}\chi'\}$ is consistent. It follows that there exists $C \in S^A$ such that $\Theta^A \cup \{\chi, {\sim}\chi'\} \subseteq C$. It follows that $\chi \in C$ and $\langle \chi', \psi' \rangle$ is not executable at C. Contradiction. Therefore, $\Theta^A \cup \{\chi, {\sim}\chi'\}$ is inconsistent. Thus, we have $\vdash \bigwedge \Theta^A \to (\chi \to \chi')$. It follows by Rule `NECU` and Axiom `DISTU` that $\vdash \mathcal{U} \bigwedge \Theta^A \to \mathcal{U}(\chi \to \chi')$. It follows by Proposition 4.4.5 that $\Theta^A \vdash \mathcal{U}(\chi \to \chi')$. $\square$

Proposition 4.4.8 *Given $\psi \in Sub^+(\varphi)$, $\langle \chi', \psi' \rangle \in Sub^+(\varphi)$ and $B \in S^A$, if $\langle \chi', \psi' \rangle$ is executable at B, and $\psi \in B'$ for each $B' \in S^A$ with $B \xrightarrow{\langle \chi', \psi' \rangle} B'$, then we have $\Theta^A \vdash \mathcal{U}(\psi' \to \psi)$.*

PROOF Assume that $\Theta^A \cup \{\psi', {\sim}\psi\}$ is consistent. It follows that there exists $C \in S^A$ such that $\Theta^A \cup \{\psi', {\sim}\psi\} \subseteq C$. It follows that $B \xrightarrow{\chi', \psi'} C$. It follows that $\psi \in C$. Contradiction. Therefore, $\Theta^A \cup \{\psi', {\sim}\psi\}$ is inconsistent. Thus, we have $\vdash \bigwedge \Theta^A \to (\psi' \to \psi)$. It follows by Rule `NECU` and Axiom `DISTU` that $\vdash \mathcal{U} \bigwedge \Theta^A \to \mathcal{U}(\psi' \to \psi)$. It follows by Proposition 4.4.5 that $\Theta^A \vdash \mathcal{U}(\psi' \to \psi)$. $\square$

Now we are ready to prove the truth lemma for the model defined in Definition 4.4.4.

Proposition 4.4.9 *For each $\psi \in Sub^+(\varphi)$, $\mathcal{M}^A, B \Vdash \psi$ iff $\psi \in B$.*

PROOF Boolean cases are trivial; we only focus on the case of $\mathcal{K}h^{\mathtt{w}}(\chi, \psi)$.

Right to Left: It follows that $\langle \chi, \psi \rangle \in Act^A$. Given $\mathcal{M}^A, C \Vdash \chi$, it follows that $\psi \in C'$ if $C \xrightarrow{\langle \chi, \psi \rangle} C'$. It follows by IH that $\mathcal{M}^A, C' \Vdash \psi$. Thus, we only need to show that $\langle \chi, \psi \rangle$ is executable at C, namely there exists $C' \in S^A$ such that $\psi \in C'$. Assume that there is no $C' \in S^A$ such that $\psi \in C'$. It follows by Proposition 4.4.6 that $\Theta^A \vdash \mathcal{U}{\sim}\psi$. Thus, we have $\Theta^A \vdash \mathcal{U}(\psi \to \bot)$. Since $\Theta^A \subseteq C$, it follows that $C \vdash \mathcal{U}(\psi \to \bot)$. Since $\mathcal{K}h^{\mathtt{w}}(\chi, \psi) \in B$, it follows that $\mathcal{K}h^{\mathtt{w}}(\chi, \psi) \in \Theta^A$, and that $\mathcal{K}h^{\mathtt{w}}(\chi, \psi) \in C$. Therefore, it follows by Axiom `UWKh` that $C \vdash \mathcal{K}h^{\mathtt{w}}(\chi, \bot)$, namely $C \vdash \mathcal{U}{\sim}\chi$. It follows by Axiom `TU` that $C \vdash {\sim}\chi$, namely $C \vdash \neg\chi$. Contradiction. Therefore, there exists $C' \in S^A$ such that $\psi \in C'$.

Left to Right: If $\mathcal{M}, B \Vdash \mathcal{K}h^{\mathtt{w}}(\chi, \psi)$, it follows that there exists $a \in (Act^A)^\bullet$ such that for each $\mathcal{M}, C \Vdash \chi$, we have a is executable at C and $\mathcal{M}, C' \Vdash \psi$ for all $C' \in S^A$ with $C \xrightarrow{a} C'$. There are three cases:

- $a = \epsilon$. It follows that $\mathcal{M}, C \Vdash \psi$ if $\mathcal{M}, C \Vdash \chi$. By IH, we have that $\chi \in C$ implies $\psi \in C$ for all $C \in S^A$. Therefore, we have $\Theta^A \cup \{\chi, \neg\psi\}$ is inconsistent. It follows that $\Theta^A \vdash \chi \to \psi$. It follows by Rule `NECU`, Axiom `DISTU`, and Proposition 4.4.5 that $\Theta^A \vdash \mathcal{U}(\chi \to \psi)$. It follows by Axiom `EMPWKh` that $\Theta^A \vdash \mathcal{K}h^{\mathtt{w}}(\chi, \psi)$. Therefore, $\mathcal{K}h^{\mathtt{w}}(\chi, \psi) \in B$.

- $a = \langle \chi', \psi' \rangle \in Act^A$ and there is no $C \in S^A$ such that $\chi \in C$. It follows that $\Theta^A \cup \{\sim\chi\}$ is inconsistent. Thus, we have $\Theta^A \vdash \neg\chi$. It follows by Rule NECU, Axiom DISTU, and Proposition 4.4.5 that $\Theta^A \vdash \mathcal{U}\neg\chi$, namely $\Theta^A \vdash \mathcal{K}h^{\mathtt{w}}(\chi, \bot)$. Since $\vdash \mathcal{U}(\bot \to \psi)$, it follows by Axiom UWKh that $\Theta^A \vdash \mathcal{K}h^{\mathtt{w}}(\chi, \psi)$. Thus, we have $\mathcal{K}h^{\mathtt{w}}(\chi, \psi) \in B$.

- $a = \langle \chi', \psi' \rangle \in Act^A$ and $\chi \in C$ for some $C \in S^A$. It follows by IH that for each $\chi \in B'$ we have $\langle \chi', \psi' \rangle$ is executable at B'. It follows by Proposition 4.4.7 that $\Theta^A \vdash \mathcal{U}(\chi \to \chi')$. It follows by IH that $\psi \in C'$ for each $C' \in S^A$ with $C \xrightarrow{\langle \chi', \psi' \rangle} C'$. It follows by Proposition 4.4.8 that $\Theta^A \vdash \mathcal{U}(\psi' \to \psi)$. Since $\langle \chi', \psi' \rangle \in Act^A$, it follows that $\Theta^A \vdash \mathcal{K}h^{\mathtt{w}}(\chi', \psi')$. It follows by Axiom UWKh that $\Theta^A \Vdash \mathcal{K}h^{\mathtt{w}}(\chi, \psi)$. Thus, we have $\mathcal{K}h^{\mathtt{w}}(\chi, \psi) \in B$.

$\square$

Proposition 4.4.10 (Completeness w.r.t. nonstandard semantics) *If* $\Vdash \varphi$ *then* $\vdash \varphi$.

PROOF We only need to show that if φ is consistent then φ is satisfiable with respect to the nonstandard semantics $\Vdash$. If φ is consistent, it follows by Proposition 4.4.3 that there is an MCS A in $Sub^+(\varphi)$ such that $\varphi \in A$. It follows by Proposition 4.4.9 that $\mathcal{M}, A \Vdash \varphi$. $\square$

Proposition 4.4.11 (Small model property) *If φ is satisfiable with respect to the nonstandard semantics, then there is a model $\mathcal{M}$ such that $\mathcal{M}, s \Vdash \varphi$ and the model is of size at most 2^k, where $k = |Sub^+(\varphi)|$.*

PROOF If φ is satisfiable with respect to the nonstandard semantics, it follows by Proposition 4.4.2 that φ is consistent. Then by Definition 4.4.4, we can construct a model $\mathcal{M}^A$ where A is an MCS in $Sub^+(\varphi)$ and $\varphi \in A$. It follows by Proposition 4.4.10 that $\mathcal{M}^A, A \Vdash \varphi$. It is obvious that $|\mathcal{M}^A| \leq O(2^{|\varphi|})$. $\square$

It follows by Propositions 4.4.2 and 4.4.10 that $\mathbb{SWKH}$ is sound and complete with respect to the nonstandard semantics $\Vdash$. Since $\mathbb{SWKH}$ is also sound and complete with respect to the standard semantics $\vDash$, we have the following lemma.

Proposition 4.4.12 $\vDash \varphi$ *if and only if* $\Vdash \varphi$.

Theorem 4.4.13 (Decidability) *The problem whether φ is valid on the standard semantics is decidable.*

PROOF To decide whether φ is valid on the standard semantics, it follows by Proposition 4.4.12 that we only need to decide whether φ is valid on the nonstandard semantics. In other words, we only need to decide whether $\neg\varphi$ is satisfiable on the nonstandard semantics. It follows by Proposition 4.4.11 that the problem whether $\neg\varphi$ is satisfiable on the nonstandard semantics is decidable. $\square$

4.5 Conclusion

In this chapter, we interpret the knowing-how formula $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi)$ as that the agent has a weak conformant plan for achieving φ given ψ, and a weak conformant plan for achieving φ-states from ψ-states is a finite linear action sequence such that the performance of the action sequence at each ψ-state will always end up with a φ-state, whether or not all parts of the plan have been completed. Our interpretation of knowledge-how is weaker than the one in Wang (2015a), where knowledge-how is interpreted as that the agent has a conformant plan, but our interpretation is more realistic. We also presented a sound and complete axiomatic system. This shows that this system is weaker than the system addressed in Wang (2015a). We also showed that this logic is decidable by reducing the problem to a decidable problem to the nonstandard semantics.

One more interesting thing is that the canonical model is more complicated than the one of Wang (2015a) even though the axiomatic system is weaker. Mainly, $\mathcal{K}h^{\mathtt{w}}$ formulas are realized by a two-step plan in our canonical model while they are realized by a one-step plan in the canonical model in Wang (2015a). This also affords us some useful ideas about how to construct the decision procedure for the logic with a tableau method. For example, for the tableau system of our logic, it is not enough to consider only one-step plans.

The nonstandard semantics played a major role in this paper not only because it is the key step in the proof of the decidability but also because it reveals the fact that our formalization of knowledge-how is in principle the same as in Moore's first interpretation. It also shows that Moore's interpretation does not contain the trivial case of knowing how to guarantee a state of affairs by doing nothing.

For future directions, we can express the existence of a weak conformant plan in the logic framework proposed in Yu et al. (2016) where the existence of a conformant plan can be expressed by a formula. Moreover, we can study the knowing-how logic under a fixed action set. In our model, the action set Act is a part of the model, but it is clear that for different Act we will get different logics. For example, if Act is empty, $\mathcal{K}h^{\mathtt{w}}(\psi, \varphi)$ is equivalent to $\mathcal{U}(\psi \to \varphi)$. If Act is a singleton, the formula $\mathcal{K}h^{\mathtt{w}}(p, q) \wedge \mathcal{K}h^{\mathtt{w}}(q, r) \to \mathcal{K}h^{\mathtt{w}}(p, r)$ will be valid under our standard semantics. The more interesting thing is to compare the logic containing a finite Act with the logic containing an infinite Act.

Another exciting research field is the multi-agent version of $\mathcal{K}h^{\mathtt{w}}$. We can also consider group notions of "knowing how". Especially, the distributed knowledge-how will be very useful. If you know how to achieve B from A and I know how to achieve C from B, we two together should know how to achieve C from A. Moreover, it also makes good sense to extend our language with a public announcement operator. The update of the new information will result in the change of the background information throughout the model, and this will affect the knowledge-how.

Chapter 5

Strategically knowing how[1]

5.1 Introduction

Standard epistemic logic focuses on reasoning about propositional knowledge expressed by *knowing that* φ (see 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 (see Fantl (2008)), e.g., whether *knowing how to swim* can be rephrased in terms of *knowing that*. In AI, it is crucial to let autonomous agents *know how* to accomplish certain goals in robotics, game playing, decision making, and multi-agent systems. 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 (see Gochet (2013)). 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 find a notion of consistency regarding knowledge databases.

Example 5.1.1 *Consider the scenario where a doctor needs a plan to treat a patient and cure his pain, under the uncertainty about some possible allergy. If there is no allergy, then simply taking some pills can cure the pain, and the surgery is not a legitimate 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. Let p denote that the patient has the pain, and let q denote that there is an allergy. The model in Figure 5.1 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 whether q*

[1] This is based on the paper Fervari et al. (2017) that is co-authored with Raul Fervari, Andreas Herzig, and Yanjing Wang.

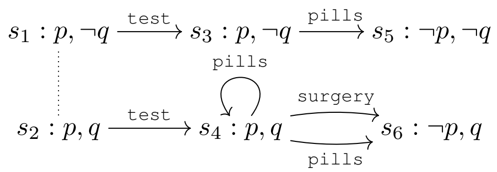


Figure 5.1: A scenario representing how to cure the pain.

happens to be positive and take the pills otherwise. We can say that the doctor in this case knows how to cure the pain.

How can we formalize the knowledge-how of the agent in such scenarios with uncertainty? In the early days of AI, people already started to look at it in the setting of logics of knowledge and action (see McCarthy and Hayes (1969); McCarthy (1979); Moore (1985); Singh (1994); 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) and simply combining the existing modalities for "knowing that" and "ability" in a logical language like ATEL (see 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 knowing that there exists a strategy to make sure p. It does not work even when we replace the notion of *strategy* by the notion of *uniform strategy* where the agent has to choose the same action on indistinguishable states (see 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* reading 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 (see Quine (1953); Stanley and Williamson (2001)). The latter 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 the *de re* one. Proposals to capture the *de re* reading have been discussed in the literature, such as making the knowledge operator more constructive (see Jamroga and Ågotnes (2007)), making the strategy explicitly specified (see Herzig et al. (2013); Belardinelli (2014)), or inserting $\mathcal{K}$ in-between an existential quantifier and the ability modality in seeing-to-it-that (STIT) logic (see Broersen and Herzig (2015)).

In Wang (2015a, 2016), a new approach is proposed by introducing a single new modality $\mathcal{K}h^{\mathsf{s}}$ 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 (2017) for a survey). The semantics of $\mathcal{K}h^{\mathsf{s}}$ is inspired by the idea of conformant planning based on linear plans (see Smith and Weld (1998); Yu et al. (2016)). It is shown that $\mathcal{K}h^{\mathsf{s}}$ 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) and Li (2017). However, in these previous works, there was no explicit knowing that modality $\mathcal{K}$ in the language and the semantics of $\mathcal{K}h^{\mathsf{s}}$ 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 5.1.1.

In this chapter, we extend this line of work in the following aspects:

- Both the *knowing how* modality $\mathcal{K}h^{\mathsf{s}}$ and *knowing that* modality $\mathcal{K}$ are in the language.
- In contrast to the state-independent semantics in Wang (2015a), we interpret $\mathcal{K}h^{\mathsf{s}}$ *locally* w.r.t. the current uncertainty.
- Instead of linear plans in Wang (2015a), the semantics of our $\mathcal{K}h^{\mathsf{s}}$ operator is based on strategies (branching plans).

The intuitive idea behind our semantics of $\mathcal{K}h^{\mathsf{s}}$ 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 (see Geffner and Bonet (2013)).

Technically, our contributions are summarized as follows:

- A logical language with both $\mathcal{K}h^{\mathsf{s}}$ 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.

This chapter is organized as follows: Section 5.2 lays out the language and semantics of our framework; Section 5.3 proposes the axiomatization and proves its soundness; we prove the completeness of our proof system and show the decidability of the logic in Section 5.4 before we conclude with future work.

5.2 The logic SKH

In this section, we will introduce our logic of knowing how with strategy, and we denote the logic as SKH. Firstly, we introduce the language of SKH. Besides the common boolean operators, there are a knowing-that modality $\mathcal{K}$ and a strategically knowing-how modality $\mathcal{K}h^{\mathsf{s}}$.

Definition 5.2.1 (Language) *Let $\mathbf{P}$ be a countable set of propositional symbols. The language is defined by the following BNF where $p \in \mathbf{P}$:*

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

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

The formula $\mathcal{K}\varphi$ reads "the agent knows that φ". The formula $\mathcal{K}h^{\mathrm{s}}\varphi$ reads "the agent strategically knows how to guarantee φ", that is, the agent has a strategy such that she knows that performing the strategy will make her know that φ.

Definition 5.2.2 (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 on W, and*
- *$V : W \to 2^{\mathbf{P}}$ is a valuation.*

Note that the labels in Act do not appear in the language. The graph in Figure 5.1.1 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 chapter we do not require any properties linking $\sim$ 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 chapter. 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. For each $s \in W$, we use $[s]$ to denote the equivalence class $\{t \in W \mid s \sim t\}$, and we 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 a $t \in W$ such that $[s] \xrightarrow{a} [t]$, we say a is executable at $[s]$.

Definition 5.2.3 (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 the knowledge-how may be trivialized. We use $\mathtt{dom}(\sigma)$ to denote the domain of σ. Function σ can be seen as a binary relation on $[W] \times Act$ such that $([s], a), ([s], b) \in \sigma$ implies $a = b$. Therefore, given strategies σ and τ with $\tau \subseteq \sigma$, it follows that $\mathtt{dom}(\tau) \subseteq \mathtt{dom}(\sigma)$, and $\tau([s]) = \sigma([s])$ for all $[s] \in \mathtt{dom}(\tau)$.

Definition 5.2.4 (Executions) *Given a strategy σ w.r.t a model $\mathcal{M}$, a* possible execution *(or just* 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 $[s] \notin \mathtt{dom}(\sigma)$. 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 $\mathtt{dom}(\sigma)$.*

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 $\mathtt{CELeaf}(\sigma, s)$ to denote the set of all leaf-nodes of all the complete executions of

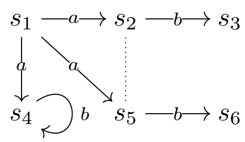


Figure 5.2

σ (which can be many due to non-determinism) starting from $[s]$, and $\mathtt{CEInner}(\sigma, s)$ to denote the set of all the inner-nodes of complete executions of σ starting from $[s]$. $\mathtt{CELeaf}(\sigma, s) \cap \mathtt{CEInner}(\sigma, s) = \emptyset$ since if $[s]$ is a leaf-node of a complete execution then σ is not defined at $[s]$.

The following is an example to help us get familiar with the notions defined above.

Example 5.2.5 *Consider the model depicted in Figure 5.2. Let σ be a function defined as $\sigma = \{\{s_1\} \mapsto a, \{s_2, s_5\} \mapsto b, \{s_4\} \mapsto b\}$. We have the following results.*

- *σ is a strategy.*
- *All the complete executions of σ starting from $[s_1]$ are:*

$$\{s_1\}\{s_2, s_5\}\{s_3\}$$
$$\{s_1\}\{s_2, s_5\}\{s_6\}$$
$$\{s_1\}\{s_4\}\{s_4\}\cdots\cdots$$

- $\mathtt{CEInner}(\sigma, s_1) = \{\{s_1\}, \{s_2, s_5\}, \{s_4\}\}$
- $\mathtt{CELeaf}(\sigma, s_1) = \{\{s_3\}, \{s_6\}\}$

Definition 5.2.6 (Semantics) *Given a pointed model $\mathcal{M}, s$, the satisfaction relation $\vDash$ is defined as follows:*

$$\begin{array}{lcl}
\mathcal{M}, s \vDash p & \Longleftrightarrow & p \in V(s) \\
\mathcal{M}, s \vDash \neg\varphi & \Longleftrightarrow & \mathcal{M}, s \nvDash \varphi \\
\mathcal{M}, s \vDash \varphi \wedge \psi & \Longleftrightarrow & \mathcal{M}, s \vDash \varphi \textit{ and } \mathcal{M}, s \vDash \psi \\
\mathcal{M}, s \vDash \mathcal{K}\varphi & \Longleftrightarrow & \textit{for all } s' : s{\sim}s' \textit{ implies } \mathcal{M}, s' \vDash \varphi \\
\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi & \Longleftrightarrow & \textit{there exists a strategy } \sigma \textit{ such that} \\
& & 1. [t] \subseteq [\![\varphi]\!] \textit{ for all } [t] \in \mathtt{CELeaf}(\sigma, s) \\
& & 2. \textit{ all its complete executions} \\
& & \quad \textit{starting from } [s] \textit{ are finite,}
\end{array}$$

where $[\![\varphi]\!] = \{s \in W \mid \mathcal{M}, s \vDash \varphi\}$.

Note that the two conditions for σ in the semantics of $\mathcal{K}h^{\mathsf{s}}$ reflect our two intuitions mentioned in the introduction. The implicit role of $\mathcal{K}$ in $\mathcal{K}h^{\mathsf{s}}$ will become more clear when the axioms are presented. Going back to Example 5.1.1, we can verify that $\mathcal{K}h^{\mathsf{s}}\neg p$ holds in s_1 and s_2 due to the strategy $\sigma = \{\{s_1, s_2\} \mapsto \mathtt{test}, \{s_3\} \mapsto \mathtt{pills}, \{s_4\} \mapsto \mathtt{surgery}\}$. Note that $\mathtt{CELeaf}(\sigma, s) = \{[s_5], [s_6]\} = \{\{s_5\}, \{s_6\}\}$ and $[\![\neg p]\!] = \{s_5, s_6\}$. On the other hand, $\mathcal{K}h^{\mathsf{s}}\neg q$ is not true in s_1: although the agent can guarantee $\neg q$ *de facto* in 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 $\mathcal{K}h^{\mathsf{s}}(p \leftrightarrow q)$ holds at s_1 and s_2 (hint: a strategy is a partial function).

The following example shows that $(\mathcal{K}h^{\mathsf{s}}\varphi \wedge \mathcal{K}h^{\mathsf{s}}(\varphi \to \psi)) \to \mathcal{K}h^{\mathsf{s}}\psi$ is not valid.

Example 5.2.7 *Let the model $\mathcal{M}$ be depicted as below. It is easy to verify that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}p$ due to the strategy $\sigma_1 = \{\{s\} \mapsto a\}$ and that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}(p \to q)$ due to the empty strategy. Since there is no strategy to get to a state where q is true, thus we have $\mathcal{M}, s \vDash \neg\mathcal{K}h^{\mathsf{s}}q$.*

$$s : \neg p, \neg q \xrightarrow{a} t : p, \neg q$$

5.3 A deductive system

In this subsection, we provide a Hilbert-style proof system for the logic SKH and show it is sound.

5.3.1 Axiom system $\mathbb{SKHS}$

Definition 5.3.1 ($\mathbb{SKH}$ System) *The axiomatic system $\mathbb{SKH}$ is shown in Table 5.1. We write $\mathbb{SKH} \vdash \varphi$ (or sometimes just $\vdash \varphi$) to mean that the formula φ is* derivable *in the axiomatic system $\mathbb{SKH}$; the negation of $\mathbb{SKH} \vdash \varphi$ is written $\mathbb{SKH} \nvdash \varphi$ (or just $\nvdash \varphi$). To say that a set D of formulas is $\mathbb{SKH}$-*inconsistent *(or just inconsistent) means that there is a finite subset $D' \subseteq D$ such that $\vdash \neg \bigwedge D'$, where $\bigwedge D' := \bigwedge_{\varphi \in D'} \varphi$ if $D' \neq \emptyset$ and $\bigwedge_{\varphi \in \emptyset} \varphi := \top$. To say that a set of formulas is $\mathbb{SKH}$-*consistent *(or just consistent) means that the set of formulas is not inconsistent. Consistency or inconsistency of a formula refers to the consistency or inconsistency of the singleton set containing the formula.*

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`	$\mathcal{K}p \to \mathcal{K}h^{\mathsf{s}}p$		
`AxKhtoKhK`	$\mathcal{K}h^{\mathsf{s}}p \to \mathcal{K}h^{\mathsf{s}}\mathcal{K}p$		
`AxKhtoKKh`	$\mathcal{K}h^{\mathsf{s}}p \to \mathcal{K}\mathcal{K}h^{\mathsf{s}}p$		
`AxKhKh`	$\mathcal{K}h^{\mathsf{s}}\mathcal{K}h^{\mathsf{s}}p \to \mathcal{K}h^{\mathsf{s}}p$		
`AxKhbot`	$\mathcal{K}h^{\mathsf{s}}\bot \to \bot$		
Rules:			
`MP`	$\dfrac{\varphi, \varphi \to \psi}{\psi}$	`NECK`	$\dfrac{\varphi}{\mathcal{K}\varphi}$
`MONOKh`	$\dfrac{\varphi \to \psi}{\mathcal{K}h^{\mathsf{s}}\varphi \to \mathcal{K}h^{\mathsf{s}}\psi}$	`SUB`	$\dfrac{\varphi(p)}{\varphi[\psi/p]}$

Table 5.1: System $\mathbb{SKHS}$

Note that we have $\mathbb{S}5$ 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 complete executions. We will come back to this axiom at the end of the chapter. 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 complete executions at all. `AxKhKh` essentially says that the strategies can be composed. Its validity is quite involved, to which we devote the next subsection. Finally, `AxKhtoKKh` is the positive introspection axiom for $\mathcal{K}h^{\mathsf{s}}$, whose validity is due to uniformity of the strategies on indistinguishable states. The corresponding negative introspection can be derived by using `AxKhtoKKh`, `5` and `T`:

Proposition 5.3.2 $\vdash \neg\mathcal{K}h^{\mathsf{s}}p \to \mathcal{K}\neg\mathcal{K}h^{\mathsf{s}}p$.

PROOF

(1)	$\neg\mathcal{K}\mathcal{K}h^{\mathsf{s}}p \to \neg\mathcal{K}h^{\mathsf{s}}p$	`AxKhtoKKh`
(2)	$\mathcal{K}\neg\mathcal{K}\mathcal{K}h^{\mathsf{s}}p \to \mathcal{K}\neg\mathcal{K}h^{\mathsf{s}}p$	$(1), \mathtt{NECK}$
(3)	$\neg\mathcal{K}\mathcal{K}h^{\mathsf{s}}p \to \mathcal{K}\neg\mathcal{K}\mathcal{K}h^{\mathsf{s}}p$	`5`
(4)	$\neg\mathcal{K}\mathcal{K}h^{\mathsf{s}}p \to \mathcal{K}\neg\mathcal{K}h^{\mathsf{s}}p$	$(2), (3), \mathtt{MP}$
(5)	$\neg\mathcal{K}h^{\mathsf{s}}p \to \neg\mathcal{K}\mathcal{K}h^{\mathsf{s}}p$	`T`
(6)	$\neg\mathcal{K}h^{\mathsf{s}}p \to \mathcal{K}\neg\mathcal{K}h^{\mathsf{s}}p$	$(4), (5), \mathtt{MP}$

□

Note that we do not have the K axiom for $\mathcal{K}h^{\mathsf{s}}$ because it is not valid (see Example 5.2.7). Instead, we have the monotonicity rule `MONOKh`. In fact, the logic is not normal, as desired, e.g., $(\mathcal{K}h^{\mathsf{s}}p \land \mathcal{K}h^{\mathsf{s}}q) \to \mathcal{K}h^{\mathsf{s}}(p \land q)$ 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.

5.3.2 Soundness

In this subsection, we show that the axiom system $\mathbb{SKH}$ is sound with respect to the semantics provided in Section 5.2. Given that the knowing-that modality $\mathcal{K}$ is interpreted in the same way as epistemic logic, we will not show that the $\mathbb{S}5$ axioms and rules for $\mathcal{K}$ are sound. Also since the logic is built on a well-understood modal logic, we will not show that the rules `MP` and `SUB` are sound. We will focus on the axioms and rules with the knowing-how modality $\mathcal{K}h^{\mathsf{s}}$.

The following proposition shows the axiom `AxKtoKh` is valid.

Proposition 5.3.3 $\vDash \mathcal{K}\varphi \to \mathcal{K}h^{\mathsf{s}}\varphi$

PROOF If $\mathcal{M}, s \vDash \mathcal{K}\varphi$, it follows that $[s] \subseteq [\![\varphi]\!]$. Let σ be the empty strategy. It follows that $[s]$ is the only complete execution of σ starting from $[s]$. Thus, we have $\mathtt{CELeaf} = \{[s]\}$. Since $[s] \subseteq [\![\varphi]\!]$, it follows that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$. □

The following proposition shows the axiom `AxKhtoKhK` is valid.

Proposition 5.3.4 $\vDash \mathcal{K}h^{\mathsf{s}}\varphi \to \mathcal{K}h^{\mathsf{s}}\mathcal{K}\varphi$

PROOF If $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$, it follows that there exists a strategy σ such that all its complete executions starting from $[s]$ are finite. Moreover, for each $[t] \in \mathtt{CELeaf}(\sigma, s)$, we have $[t] \subseteq [\![\varphi]\!]$, and then we have $\mathcal{M}, t' \vDash \mathcal{K}\varphi$ for all $t' \in [t]$. Therefore, we have $[t] \subseteq [\![\mathcal{K}\varphi]\!]$ for each $[t] \in \texttt{CELeaf}(\sigma, s)$. Thus, we have $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\mathcal{K}\varphi$. □

The following proposition shows that the axiom `AxKhtoKKh` is valid.

Proposition 5.3.5 $\vDash \mathcal{K}h^{\mathsf{s}}\varphi \to \mathcal{K}\mathcal{K}h^{\mathsf{s}}\varphi$

PROOF If $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$, it follows that there exists a strategy σ such that all its complete executions starting from $[s]$ are finite. Moreover, for each $[t] \in \texttt{CELeaf}(\sigma, s)$, we have $[t] \subseteq [\![\varphi]\!]$. Given $s' \in [s]$, since $[s] = [s']$, it follows that all complete executions of σ starting from $[s']$ are the same as all complete executions of σ starting from $[s]$. Thus, we have $\texttt{CELeaf}(\sigma, s) = \texttt{CELeaf}(\sigma, s')$. Therefore, we have $\mathcal{M}, s' \vDash \mathcal{K}h^{\mathsf{s}}\varphi$ for all $s' \in [s]$. It follows that $\mathcal{M}, s \vDash \mathcal{K}\mathcal{K}h^{\mathsf{s}}\varphi$. □

The following proposition shows that the axiom `AxKhbot` is valid.

Proposition 5.3.6 $\vDash \mathcal{K}h^{\mathsf{s}}\bot \to \bot$

PROOF We only need to show $\vDash \neg\mathcal{K}h^{\mathsf{s}}\bot$. Assuming that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\bot$ for some pointed model $\mathcal{M}, s$. It follows that there exists a strategy σ such that all its complete executions starting from $[s]$ are finite and $[t] \subseteq [\![\bot]\!]$ for each $[t] \in \texttt{CELeaf}(\sigma, s)$. No matter whether $[s] \in \texttt{dom}(\sigma)$ or not, we always have $\texttt{CELeaf}(\sigma, s) \neq \emptyset$. It follows that there exists $[t] \in \texttt{CELeaf}(\sigma, s)$ such that $[t] \subseteq [\![\bot]\!]$. Since $[t] \neq \emptyset$ and $[\![\bot]\!] = \emptyset$, it is a contradiction that $[t] \subseteq [\![\bot]\!]$. Therefore, $\mathcal{K}h^{\mathsf{s}}\bot$ is not satisfiable, and thus $\vDash \neg\mathcal{K}h^{\mathsf{s}}\bot$. □

The following proposition shows that the rule `MONOKh` preserves the validity.

Proposition 5.3.7 *If* $\vDash \varphi \to \psi$ *then we have* $\vDash \mathcal{K}h^{\mathsf{s}}\varphi \to \mathcal{K}h^{\mathsf{s}}\psi$.

PROOF If $\vDash \varphi \to \psi$ and $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$, we need to show $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\psi$. It follows by $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$ that there exists a strategy σ such that all its complete executions starting from $[s]$ are finite. Moreover, for each $[t] \in \texttt{CELeaf}(\sigma, s)$, we have $[t] \subseteq [\![\varphi]\!]$. Since $\vDash \varphi \to \psi$, it follows that $[\![\varphi]\!] \subseteq [\![\psi]\!]$. Thus, we have $[t] \subseteq [\![\psi]\!]$ for each $[t] \in \texttt{CELeaf}(\sigma, s)$. Therefore, we have $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\psi$. □

The axiom `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 ($\mathcal{K}h^{\mathsf{s}}\mathcal{K}h^{\mathsf{s}}p$). 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 ($\mathcal{K}h^{\mathsf{s}}p$). The general idea is actually simple: first order those leafnodes $[s]$ (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 5.3.8 *Given strategies* τ *and* σ *with* $\tau \subseteq \sigma$, *if* $[s] \in \texttt{dom}(\tau)$ *and* $\texttt{dom}(\sigma) \cap \texttt{CELeaf}(\tau, s) = \emptyset$, *then a sequence is a complete execution of* σ *from* $[s]$ *if and only if it is a complete execution of* τ *from* $[s]$.

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

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

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

Now, we are ready to show that the axiom $\texttt{AxKhKh}$ is valid.

Proposition 5.3.9 *$\vDash \mathcal{K}h^{\mathsf{s}}\mathcal{K}h^{\mathsf{s}}\varphi \to \mathcal{K}h^{\mathsf{s}}\varphi$.*

PROOF Supposing $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\mathcal{K}h^{\mathsf{s}}\varphi$, we will show that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$. It follows by the semantics that there exists a strategy σ such that all complete executions of σ from $[s]$ are finite and $[t] \subseteq [\![\mathcal{K}h^{\mathsf{s}}\varphi]\!]$ for all $[t] \in \texttt{CELeaf}(\sigma, s)$ $(*)$. If $[s] \notin \texttt{dom}(\sigma)$, then $\texttt{CELeaf}(\sigma, s) = \{[s]\}$, and then it is trivial that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$. Next we focus on the case of $[s] \in \texttt{dom}(\sigma)$.

According to the well-ordering theorem (equivalent to Axiom of Choice), we assume $\texttt{CELeaf}(\sigma, s) = \{S_i \mid i < \gamma\}$ 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 $\mathcal{M}, s_i \vDash \mathcal{K}h^{\mathsf{s}}\varphi$ for each $i < \gamma$, it follows that for each $[s_i]$ there exists a strategy σ_i such that all complete executions of σ_i from $[s_i]$ are finite and $[v] \subseteq [\![\varphi]\!]$ for all $[v] \in \texttt{CELeaf}(\sigma_i, s_i)$ $(\blacktriangleleft)$. Next, in order to show $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$, 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 $D_i = \texttt{CEInner}(\sigma_i, s_i) \setminus (\texttt{dom}(f_i) \cup \{[v] \in \texttt{CELeaf}(f_i, t) \mid [t] \in \texttt{dom}(f_i)\})$ we define:

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

Claim 5.3.9.1 *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 $0 \leq i < \gamma$, $\texttt{dom}(\tau_i) \cap \texttt{CELeaf}(\tau_j, t) = \emptyset$ where $t \in \texttt{dom}(\tau_j)$ if $j < i$;*
4. *For each $0 \leq i < \gamma$, if $\delta = [t_0]\cdots$ is a complete execution of τ_i from $[t] \in \texttt{dom}(\tau_i)$ then $|\delta| = n$ for some $n \in \mathbb{N}$ and $[t_n] \subseteq [\![\varphi]\!]$;*

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

Proof of claim 5.3.9.1:

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 $\mathtt{dom}(f_\alpha) \cap D_\alpha = \emptyset$. Since $D_\alpha = \mathtt{CEInner}(\sigma_\alpha, s_\alpha) \setminus \mathtt{dom}(f_\alpha) \setminus \{[v] \in \mathtt{CELeaf}(f_\alpha, t) \mid t \in \mathtt{dom}(f_\alpha)\}$, it is obvious that $\mathtt{dom}(f_\alpha) \cap D_\alpha = \emptyset$.

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 $t \in \mathtt{dom}(\tau_j)$, we need to show that $\mathtt{dom}(\tau_\alpha) \cap \mathtt{CELeaf}(\tau_j, t) = \emptyset$. Supposing $[v] \in \mathtt{CELeaf}(\tau_j, t)$, we will show that $[v] \notin \mathtt{dom}(\tau_\alpha)$, namely $[v] \notin \mathtt{dom}(f_\alpha) \cup D_\alpha$. Since $j < \alpha$ and $f_\alpha = \bigcup_{\beta<\alpha} \tau_\alpha$, it follows $t \in \mathtt{dom}(f_\alpha)$. Moreover, due to $D_\alpha = \mathtt{CEInner}(\sigma_\alpha, s_\alpha) \setminus \mathtt{dom}(f_\alpha) \setminus \{[v] \in \mathtt{CELeaf}(f_\alpha, t) \mid t \in \mathtt{dom}(f_\alpha)\}$, it follows $[v] \notin D_\alpha$.

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

4. We prove it by induction on i. For the case of $i = 0$, due to the fact that $\mathtt{dom}(\tau_0) = \mathtt{CEInner}(\sigma_0, s_0)$, 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 complete execution of τ_0 from $[t]$, it follows that μ is a complete execution of σ_0 from $[s_0]$. It follows by (◀) that μ is finite. Thus, $\delta = [t_0] \cdots [t_n]$ for some $n \in \mathbb{N}$. Since $[t_n] \in \mathtt{CELeaf}(\sigma_0, s_0)$, it follows by (◀) that $[t_n] \subseteq [\![\varphi]\!]$.

   For the case of $i = \alpha > 0$, there are two situations: $[t] \in \mathtt{dom}(f_\alpha)$ or $[t] \in D_\alpha$. If $[t] \in \mathtt{dom}(f_\alpha)$, it follows that $[t] \in \mathtt{dom}(\tau_\beta)$ for some $\beta < \alpha$. By 3, we have $\mathtt{dom}(\tau_\alpha) \cap \mathtt{CELeaf}(\tau_\beta, t) = \emptyset$. Since δ is a complete execution of τ_α, it follows by Proposition 5.3.8 that δ is also a complete execution of τ_β 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. $[t_k] \in \mathtt{dom}(\tau_\beta)$, or there do not exist such k and β. (Please note that $|\delta| > 1$ due to the fact that $\delta = [t_0] \cdots$ is a complete execution of τ_α from $[t] \in \mathtt{dom}(\tau_\alpha)$).

   - $[t_k] \in \mathtt{dom}(\tau_\beta)$ for some $k < |\delta|$ and some $\beta < \alpha$: It follows that $\mu = [t_k] \cdots$ is a complete execution of τ_α from $[t_k]$. By 3. and Proposition 5.3.8, μ is a complete execution of τ_β 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 do not exist $k < |\delta|$ and $\beta < \alpha$ s.t. $[t_k] \in \mathtt{dom}(\tau_\beta)$, it follows that $\delta = [t_0] \cdots$ is a σ_α's possible execution from $[t]$. Since $[t] \in D_\alpha \subseteq$

$\texttt{CEInner}(\sigma_\alpha, s_\alpha)$, then 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 complete executions of σ_α 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 complete execution of τ_α from t and it is also a possible execution of σ_α from t, there are two cases: $[t_n] \in \texttt{CELeaf}(f_\alpha, t')$ for some $t' \in \texttt{dom}(f_\alpha)$, or δ is a complete execution of σ_α from t. If $[t_n] \in \texttt{CELeaf}(f_\alpha, t')$ for some $t' \in \texttt{dom}(f_\alpha)$, then there exists $\beta < \alpha$ s.t. $[t] \in \texttt{CELeaf}(\tau_\beta, t')$ and $[t'] \in \texttt{dom}(\beta)$. By IH, $[t_n] \subseteq [\![\varphi]\!]$. If δ is a complete execution of σ_α from t, it follows that μ is a complete execution of σ_α from $[s_\alpha]$. Then by ($\blacktriangleleft$), we have $[t_n] \subseteq [\![\varphi]\!]$.

5. If $[s_i] \notin \texttt{dom}(\sigma_i)$, 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 $[s_0] \in \texttt{dom}(\sigma_0)$ that $[s_0] \in \texttt{CEInner}(\sigma_0, s_0)$. Thus, $[s_0] \in \texttt{dom}(\tau_0)$.

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

■

Step II. We define $\tau_\gamma = \bigcup_{i<\gamma} \tau_i$. It follows by 1. and 2. of Claim 5.3.9.1 that τ_γ is indeed a partial function. Then we prove the following claim.

Claim 5.3.9.2 *If $\delta = [t_0]\cdots$ is a complete execution of τ_γ from $[t] \in \texttt{dom}(\tau_\gamma)$ then $|\delta| = n$ for some $n \in \mathbb{N}$ and $[t_n] \subseteq [\![\varphi]\!]$*

Proof of claim 5.3.9.2: Since $[t] \in \texttt{dom}(\tau_\gamma)$, it follows that $[t] \in \texttt{dom}(\tau_i)$ for some $i < \gamma$. It follows by 5. of Claim 5.3.9.1 that all complete executions of τ_i from $[t]$ are finite. Thus, there exists $\mu \sqsubseteq \delta$ such that $|\mu| = n$ for some $n \in \mathbb{N}$ and μ is a complete execution of τ_i from $[t]$. It follows by 5. of Claim 5.3.9.1 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 $\{t_k \mid 0 \leq k \leq n\} \subseteq \texttt{dom}(\tau_j)$. It cannot be that $j \leq i$. Otherwise, μ is not a complete execution of τ_i since $\tau_j \subseteq \tau_i$ by 1. of Claim 5.3.9.1. Thus, we have $j > i$. Since we also have that $[t_n] \in \texttt{dom}(\tau_j)$, $[t_n] \in \texttt{CELeaf}(\tau_i, t)$ and $t \in \texttt{dom}(\tau_i)$, this is contradictory with 3. of Claim 5.3.9.1. Therefore, we have $\delta = \mu$. ■

Step III. We define τ as $\tau = \tau_\gamma \cup (\sigma|_C)$ where $C = \texttt{CEInner}(\sigma, s) \setminus (\texttt{dom}(\tau_\gamma) \cup \{[v] \in \texttt{CELeaf}(\tau', t) \mid [t] \in \texttt{dom}(\tau_\gamma)\})$ 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 5.3.9.3 *If $\delta = [t_0]\cdots$ is a complete execution of τ from $[t] \in \texttt{dom}(\tau)$ then $|\delta| = n$ for some $n \in \mathbb{N}$ and $[t_n] \subseteq [\![\varphi]\!]$.*

Proof of claim 5.3.9.3: Since $\mathtt{dom}(\tau) = \mathtt{dom}(\tau_\gamma) \cup C$, there are two cases: $[t] \in \mathtt{dom}(\tau_\gamma)$ or $[t] \in C$.

If $[t] \in \mathtt{dom}(\tau_\gamma)$, it follows that $\mathtt{CELeaf}(\tau_\gamma, t) \cap C = \emptyset$. Moreover, we have $\mathtt{CELeaf}(\tau_\gamma, t) \cap \mathtt{dom}(\tau_\gamma) = \emptyset$. Thus, we have $\mathtt{CELeaf}(\tau_\gamma, t) \cap \mathtt{dom}(\tau) = \emptyset$. It follows by Proposition 5.3.8 that δ is a complete execution of τ_γ from from $[t]$. It follows by Claim 5.3.9.2 $|\delta| = n$ 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 $[t_k] \in \mathtt{dom}(\tau_\gamma)$, or there does not exists such k. (Please note that $|\delta| > 1$ due to the fact that $\delta = [t_0] \cdots$ is complete execution of τ from $[t] \in \mathtt{dom}(\tau)$).

- $[t_k] \in \mathtt{dom}(\tau_\gamma)$ for some $k < |\delta|$: It follows that $\mu = [t_k] \cdots$ is a complete execution of τ from $[t_k]$. Since $\mathtt{dom}(\tau) \cap \mathtt{CELeaf}(\tau_\gamma, t_k) = \emptyset$. It follows by Proposition 5.3.8 that μ is a complete execution of τ_γ from $[t_k]$. It follows by Claim 5.3.9.2 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. $[t_k] \in \mathtt{dom}(\tau_\gamma)$, then $\delta = [t_0] \cdots$ is a σ's possible execution from $[t]$. Since $[t] \in C \subseteq \mathtt{CEInner}(\sigma, s)$, 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 complete executions of σ 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 complete execution of τ from t and it is also a σ's possible execution from t, there are two cases: $[t_n] \in \mathtt{CELeaf}(\tau_\gamma, t')$ for some $t' \in \mathtt{dom}(\tau_\gamma)$, or δ is a complete execution of σ from t. If $[t_n] \in \mathtt{CELeaf}(\tau_\gamma, t')$ for some $[t'] \in \mathtt{dom}(\tau_\gamma)$, it follows by Claim 5.3.9.2 that $[t_n] \subseteq [\![\varphi]\!]$. If δ is a complete execution of σ from t, it follows that μ is a complete execution of σ from $[s]$. It follows that $[t_n] = S_i$ for some $0 \leq i < \gamma$. Since $\delta = [t_0] \cdots [t_n]$ is a complete execution of τ from $[t] \in \mathtt{dom}(\tau_\gamma)$, it follows $[t_n] \notin \mathtt{dom}(\tau_\gamma)$. We then have $[t_n] \notin \mathtt{dom}(\tau_i)$, namely $S_i \notin \tau_i$. It follows by 5. of Claim 5.3.9.1 that $S_i \subseteq [\![\varphi]\!]$, namely $[t_n] \subseteq [\![\varphi]\!]$.

■

Next, we continue to show that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$ with the assumption that $[s] \in \mathtt{dom}(\sigma)$. Since $[s] \in \mathtt{dom}(\sigma)$, we have $[s] \in \mathtt{CEInner}(\sigma, s)$. There are two cases: $[s] \in \mathtt{dom}(\tau)$ or not. If $[s] \in \mathtt{dom}(\tau)$, it follows by claim 5.3.9.3 that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$. If $[s] \notin \mathtt{dom}(\tau)$, due to $[s] \in \mathtt{CEInner}(\sigma, s)$, it follows that $[s] \in \mathtt{CELeaf}(\tau_\gamma, t)$ for some $[t] \in \mathtt{dom}(\tau_\gamma)$. It follows by Claim 5.3.9.2 that $[s] \subseteq [\![\varphi]\!]$. It follows that $\mathcal{M}, s \vDash \mathcal{K}\varphi$. It is obvious that $\mathcal{M}, s \vDash \mathcal{K}h^{\mathsf{s}}\varphi$. □

Now we are ready to prove the soundness, which can be proven by induction on the length of the proof.

Theorem 5.3.10 (Soundness) *If* $\vdash \varphi$ *then* $\vDash \varphi$.

5.4 Completeness and decidability

Let Φ be a set of formulas such that it is closed under subformula and $\sim\varphi \in \Phi$ for each $\varphi \in \Phi$, where $\sim\varphi = \chi$ if $\varphi = \neg\chi$, otherwise, $\sim\varphi = \neg\varphi$. It is obvious that Φ is countable since the whole language itself is countable. Given a set of formulas Δ, let

$$\begin{aligned}
\Delta|_{\mathcal{K}} &= \{\mathcal{K}\varphi \mid \mathcal{K}\varphi \in \Delta\} \\
\Delta|_{\neg\mathcal{K}} &= \{\neg\mathcal{K}\varphi \mid \neg\mathcal{K}\varphi \in \Delta\} \\
\Delta|_{\mathcal{K}h^{\mathsf{s}}} &= \{\mathcal{K}h^{\mathsf{s}}\varphi \mid \mathcal{K}h^{\mathsf{s}}\varphi \in \Delta\} \\
\Delta|_{\neg\mathcal{K}h^{\mathsf{s}}} &= \{\neg\mathcal{K}h^{\mathsf{s}}\varphi \mid \neg\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta\}
\end{aligned}$$

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 5.4.1 *$cl(\Phi)$ is the smallest set such that:*

- $\Phi \subseteq cl(\Phi)$*;*
- *if $\varphi \in \Phi$ then $\mathcal{K}\varphi, \neg\mathcal{K}\varphi \in cl(\Phi)$.*

Definition 5.4.2 (Atom) *Let $cl(\Phi) = \{\psi_i \mid i \in \mathbb{N}\}$. The formula set $\Delta = \{Y_i \mid i \in \mathbb{N}\}$ is an atom of $cl(\Phi)$ if*

- *$Y_i = \psi_i$ or $Y_i = {\sim}\psi_i$ for all $\psi_i \in cl(\Phi)$;*
- *Δ is consistent.*

An atom of Φ is also called an maximally consistent subset of Φ. Note that if Φ is the whole language then an atom is simply a maximally consistent set. By a standard inductive construction, we can obtain the Lindenbaum-like result in our setting:

Proposition 5.4.3 *Let Γ be a consistent subset of $cl(\Phi)$ 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 $\pm\varphi = {\sim}\varphi$.*

PROOF Let $\psi_1, \cdots, \psi_n, \cdots$ be all the formulas in $cl(\Phi) \setminus \Gamma \setminus \{\varphi\}$. We define Γ_i as below.

$$\Gamma_0 = \Gamma \cup \{\pm\varphi\}$$
$$\Gamma_{i+1} = \begin{cases} \Gamma_i \cup \{\psi_i\} & \text{if } \Gamma_i \cup \{\psi_i\} \text{ is consistent} \\ \Gamma_i \cup \{{\sim}\psi_i\} & \text{otherwise} \end{cases}$$

Firstly, we will show that Γ_i is consistent for all $i \in \mathbb{N}$. Since Γ_0 is consistent, we only need to show that if Γ_i is consistent then Γ_{i+1} is consistent, i.e. either $\Gamma_i \cup \{\psi_i\}$ or $\Gamma_i \cup \{{\sim}\psi_i\}$ is consistent. Assuming both $\Gamma_i \cup \{\psi_i\}$ and $\Gamma_i \cup \{\neg\psi_i\}$ are not consistent, it follows that $\Gamma_i \vdash \neg\psi_i$ and $\Gamma_i \vdash \psi_i$. That is, Γ_i is inconsistent. Contradiction. Therefore, either $\Gamma_i \cup \{\psi_i\}$ or $\Gamma_i \cup \{{\sim}\psi_i\}$ is consistent.

Let $\Delta' = \bigcup_{i \in \mathbb{N}} \Gamma_i$. It follows that Δ' is consistent. It is obvious that either $\psi \in \Delta'$ or ${\sim}\psi \in \Delta'$ for all $\psi \in cl(\Phi)$. Therefore, Δ' is an atom of $cl(\Phi)$. □

Next, we are going to build a canonical model w.r.t. Φ.

Definition 5.4.4 *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* $= \{\varphi \mid \mathcal{K}h^{\mathsf{s}}\varphi \in \Phi\}$;
- $\Delta \sim \Delta'$ *iff* $\Delta|_{\mathcal{K}} = \Delta'|_{\mathcal{K}}$;
- *for each* $\varphi \in$ *Act*, $\Delta \xrightarrow{\varphi} \Delta'$ *iff* $\mathcal{K}h^{\mathsf{s}}\varphi, \neg\mathcal{K}\varphi \in \Delta$ *and* $\mathcal{K}\varphi \in \Delta'$;
- *for each* $p \in \Phi$, $p \in V(\Delta)$ *iff* $p \in \Delta$.

Note that we use formulas that the agent knows how to achieve as the action labels, and we introduce an action transition if it is necessary, i.e., $\mathcal{K}h^{\mathsf{s}}\varphi$ but $\neg\mathcal{K}\varphi$ (the empty strategy does not work). Requiring $\mathcal{K}\varphi \in \Delta'$ is to reflect the first condition in the semantics of $\mathcal{K}h^{\mathsf{s}}$. Using `NECK`, `DISTK` and Proposition 5.4.3, it is routine to show the existence lemma for $\mathcal{K}$:

Proposition 5.4.5 *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* $\sim\varphi \in \Delta'$.

PROOF Let $\Gamma = \Delta|_{\mathcal{K}} \cup \Delta|_{\neg\mathcal{K}} \cup \{\sim\varphi\}$. 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 such that

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

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 $1 \leq i \leq n$ and $\vdash \mathcal{K}\neg\mathcal{K}\psi_i \leftrightarrow \neg\mathcal{K}\psi_i$ for all $1 \leq i \leq m$, 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 5.4.3 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, $\Delta' \in [\Delta]$. □

Proposition 5.4.6 *Let* Δ *and* Δ' *be two states in* $\mathcal{M}^{\Phi}$ *such that* $\Delta \sim \Delta'$. *We have* $\Delta|_{\mathcal{K}h^{\mathsf{s}}} = \Delta'|_{\mathcal{K}h^{\mathsf{s}}}$.

PROOF For each $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$, by Definition 5.4.1, $\mathcal{K}h^{\mathsf{s}}\varphi \in \Phi$, and then $\mathcal{K}\mathcal{K}h^{\mathsf{s}}\varphi \in cl(\Phi)$. For each $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$, by Axiom `AxKhtoKKh`, we have $\mathcal{K}\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$. Since $\Delta \sim \Delta'$, then $\mathcal{K}\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$, and by Axiom `T`, $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$. So we showed that $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$ implies $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$. Similarly we can prove $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$ implies $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$. Hence, $\Delta|_{\mathcal{K}h^{\mathsf{s}}} = \Delta'|_{\mathcal{K}h^{\mathsf{s}}}$. □

The following is a crucial observation for proofs.

Proposition 5.4.7 *Let* Δ *be a state in* $\mathcal{M}^{\Phi}$ *and* $\psi \in$ *Act be executable at* $[\Delta]$. *If* $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$ *for all* Δ' *with* $[\Delta] \xrightarrow{\psi} [\Delta']$ *then* $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$.

PROOF First, we show that $\mathcal{K}\psi$ is not consistent with $\neg\mathcal{K}h^{\mathsf{s}}\varphi$. It is obvious that $\mathcal{K}h^{\mathsf{s}}\varphi \in cl(\Phi)$. 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 $\neg\mathcal{K}h^{\mathsf{s}}\varphi$, by Proposition 5.4.3 there exists an atom Γ of $cl(\Phi)$ s.t. $\{\mathcal{K}\psi, \neg\mathcal{K}h^{\mathsf{s}}\varphi\} \subseteq \Gamma$. Since $\psi \in Act$ is executable at $[\Delta]$, then by definition of $\xrightarrow{\psi}$, $\sim$ and Proposition 5.4.6, $\mathcal{K}h^{\mathsf{s}}\psi, \neg\mathcal{K}\psi \in \Delta$. It follows that $\Delta \xrightarrow{\psi} \Gamma$, then $[\Delta] \xrightarrow{\psi} [\Gamma]$. This is contradictory with the assumption that $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$ for all Δ' with $[\Delta] \xrightarrow{\psi} [\Delta']$. Then $\mathcal{K}\psi$ is not consistent with $\neg\mathcal{K}h^{\mathsf{s}}\varphi$. Hence, $\vdash \mathcal{K}\psi \to \mathcal{K}h^{\mathsf{s}}\varphi$.

Since $\vdash \mathcal{K}\psi \to \mathcal{K}h^{\mathsf{s}}\varphi$, it follows by Rule `MONOKh` and Axiom `AxKhtoKhK` that $\vdash \mathcal{K}h^{\mathsf{s}}\psi \to \mathcal{K}h^{\mathsf{s}}\mathcal{K}h^{\mathsf{s}}\varphi$. Moreover, it follows by Axiom `AxKhKh` that $\vdash \mathcal{K}h^{\mathsf{s}}\psi \to \mathcal{K}h^{\mathsf{s}}\varphi$. Since ψ is executable at $[\Delta]$, it follows by the definition of $\xrightarrow{\psi}$ and Proposition 5.4.6 that $\mathcal{K}h^{\mathsf{s}}\psi \in \Delta$. Therefore, we have $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$. $\square$

Lemma 5.4.8 *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 cases of $\mathcal{K}\varphi$ and $\mathcal{K}h^{\mathsf{s}}\varphi$; the other cases are straightforward.

- Case of $\mathcal{K}\varphi$. If $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}\varphi$, we will show $\mathcal{K}\varphi \in \Delta$. Assuming $\mathcal{K}\varphi \notin \Delta$, it follows by Proposition 5.4.5 that there exists $\Delta' \in [\Delta]$ such that $\neg\varphi \in \Delta'$. It follows by IH that $\mathcal{M}^{\Phi}, \Delta' \vDash \neg\varphi$. It is contradictory with $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}\varphi$ and $\Delta' \in [\Delta]$. Thus we have $\mathcal{K}\varphi \in \Delta$.

  If $\mathcal{K}\varphi \in \Delta$, we will show $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}\varphi$. Assuming $\mathcal{M}^{\Phi}, \Delta \nvDash \mathcal{K}\varphi$, it follows that there is $\Delta' \in [\Delta]$ such that $\mathcal{M}^{\Phi}, \Delta' \vDash \neg\varphi$. It follows by IH that $\neg\varphi \in \Delta'$. It must be the case of $\neg\mathcal{K}\varphi \in \Delta'$ because $\mathcal{K}\varphi \in \Delta'$ implies $\varphi \in \Delta'$. It follows that $\neg\mathcal{K}\varphi \in \Delta$. Contradiction. Thus we have $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}\varphi$.

- Case of $\mathcal{K}h^{\mathsf{s}}\varphi$. Note that if $\mathcal{K}h^{\mathsf{s}}\varphi \in cl(\Phi)$ then $\varphi \in cl(\Phi)$ thus by Definition 5.4.1 $\mathcal{K}\varphi \in cl(\Phi)$.

  Right to Left: If $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$, we will show $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}h^{\mathsf{s}}\varphi$. 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 $\mathcal{M}^{\Phi}, \Delta \vDash \mathcal{K}h^{\mathsf{s}}\varphi$. 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 $\vdash \mathcal{K}h^{\mathsf{s}}\mathcal{K}\varphi \to \mathcal{K}h^{\mathsf{s}}\bot$. It follows by Axiom `AxKhbot` that $\vdash \mathcal{K}h^{\mathsf{s}}\mathcal{K}\varphi \to \bot$. Since $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$, it follows by Axiom `AxKhtoKhK` that $\Delta \vdash \bot$, which is contradictory with the fact that Δ is consistent. Therefore, $\mathcal{K}\varphi$ is consistent.

  By Proposition 5.4.3 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 $[\Delta''] \subseteq [\![\varphi]\!]$. 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 $[\Delta''] \subseteq [\![\varphi]\!]$ for each $[\Delta''] \in \texttt{CELeaf}(\sigma, \Delta)$. Then, $\mathcal{M}^\Phi, \Delta \vDash \mathcal{K}h^{\mathsf{s}}\varphi$.

Left to Right: Suppose $\mathcal{M}^\Phi, \Delta \vDash \mathcal{K}h^{\mathsf{s}}\varphi$, we will show $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$. By the semantics, there exists a strategy σ s.t. all complete executions of σ starting from $[\Delta]$ are finite and $[\Gamma] \subseteq [\![\varphi]\!]$ 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 5.4.5, we get $\mathcal{K}\varphi \in \Gamma$ for all $[\Gamma] \in \texttt{CELeaf}(\sigma, \Delta)$. By Axiom $\texttt{AxKtoKh}$ and Proposition 5.4.6, $\mathcal{K}h^{\mathsf{s}}\varphi \in \Gamma$ for all $[\Gamma] \in \texttt{CELeaf}(\sigma, \Delta)$.

If $[\Delta] \notin \texttt{dom}(\sigma)$, it is obvious that $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$ 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 $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$, we will show a more strong result that $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$ for all $[\Delta'] \in \texttt{CEInner}(\sigma, \Delta)$. Firstly, we show the following claim:

Claim 5.4.8.1 *If there exists* $[\Delta'] \in \texttt{CEInner}(\sigma, \Delta)$ *such that* $\neg\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$ *then there exists an infinite execution of* σ *starting from* $[\Delta]$.

Proof of claim 5.4.8.1: Let X be the set $\{[\Theta] \in \texttt{CEInner}(\sigma, \Delta) \mid \neg\mathcal{K}h^{\mathsf{s}}\varphi \in \Theta\}$. 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 $\neg\mathcal{K}h^{\mathsf{s}}\varphi \in \Theta$, by Proposition 5.4.7 there exists an atom Θ' s.t. $[\Theta] \xrightarrow{\sigma([\Theta])} [\Theta']$ and $\neg\mathcal{K}h^{\mathsf{s}}\varphi \in \Theta'$. Since $\mathcal{K}h^{\mathsf{s}}\varphi \in \Gamma$ 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 $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$ for all $[\Delta'] \in \texttt{CEInner}(\sigma, s)$. Otherwise, by claim 5.4.8.1 there is an infinite complete execution given σ from $[\Delta]$. This is contradictory with the fact that if all complete execution of σ from $[\Delta]$ are finite, then $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta'$ for all $[\Delta'] \in \texttt{CEInner}(\sigma, s)$. Since $[\Delta] \in \texttt{dom}(\sigma)$, we get $[\Delta] \in \texttt{CEInner}(\sigma, \Delta)$. Then $\mathcal{K}h^{\mathsf{s}}\varphi \in \Delta$.

□

Please note that in the proofs above it does not matter whether the domain of $\mathcal{M}^\Phi$ is finite or not. Therefore, let Φ be the set of all formulas, then each maximally 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 5.4.9 $\mathbb{SKH}$ *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 5.4.3, 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, so decidability follows.

Proposition 5.4.10 (Small model property) *If φ_0 is satisfiable then it is satisfiable in a model with at most 2^k states where $k = |cl(\Phi)|$ and Φ is the subformula closure generated by φ_0.*

Theorem 5.4.11 *SKH is decidable.*

PROOF With the small model property, this can be proved by a standard argument presented in Blackburn et al. (2001). □

5.5 Conclusion

In this chapter, we propose an epistemic logic of both (goal-directed) knowing how and knowing that, and capture the interaction of the two. We have shown that this logic is sound and complete, as well as decidable. We hope that 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* ($\mathcal{K}[a]\varphi \to [a]\mathcal{K}\varphi$) and/or *no learning* ($[a]\mathcal{K}\varphi \to \mathcal{K}[a]\varphi$) (cf. e.g., Fagin et al. (1995); Wang and Li (2012)) can change the logic or not.

Our notion of knowing how is relatively strong, particularly evidenced by the axiom `AxKhtoKhK` : $\mathcal{K}h^{\mathsf{s}}\varphi \to \mathcal{K}h^{\mathsf{s}}\mathcal{K}\varphi$, which is due to the first condition of our semantics for $\mathcal{K}h^{\mathsf{s}}$, inspired by planning with uncertainty. We believe that this 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. Yu et al. (2016)). 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. Another obvious next step is to consider knowing how in multi-agent settings.

Chapter 6

Privacy in arrow update logic[1]

6.1 Introduction

Information plays an important role in several fields of scientific research, such as philosophy, game theory, and artificial intelligence. In this chapter, the notion of information is confined to the kind of information in one's mind, which can also be called *belief* or *knowledge*. In real-life contexts, information is often communicated. This leads to a change of agents' information without any change in the bare facts of the world. One kind of these communicative events is *announcement*. This chapter will focus on reasoning about information change due to announcements.

In a multi-agent system, there are at least three types of announcements: public, private and semi-private (cf. e.g., Baltag and Moss (2004)). Imagine a scenario where two agents a and b are in a room, and in front of them, there is a coin in a closed box. Neither of them knows whether the coin is lying heads up or tails up. A public announcement occurs when the box is opened for both to see. This changes not only the agent's information about the bare facts (basic information) but also agents' information about each other (higher-order information). When a secretly opens the box and b does not suspect that anything happened, the effect is the same as the effect that the truth is privately announced to a. This changes only a's basic and higher-order information. A semi-private announcement occurs when a opens the box and b observes a's action but b does not see the coin. This changes a's basic information and the higher-order information of both.

There are a great number of modal logic theories which formalize reasoning about information change. Plaza's logic (see Plaza (1989, 2007)) is concerned with reasoning about information change due to public announcement, in which a public announcement of a statement eliminates all epistemic possibilities in which the statement does not hold. Gerbrandy and Groeneveld develop a more general dynamic epistemic logic in Gerbrandy and Groeneveld (1997), which formalizes reasoning about information change produced by public and private announcements. The dynamic epistemic logic with action models (DEL), due to Baltag et al. (1998) (see also Baltag and Moss (2004); van Ditmarsch et al. (2007)), is a powerful tool to formalize reasoning about information change. An action model is a Kripke model-like object that describes agents' beliefs

[1]This is based on a short paper presented on the conference of *Advances in Modal Logic 2014*.

about incoming information. All public, private and semi-private announcement can be modeled in DEL.

Kooi and Renne's Arrow Update Logic (AUL) (see Kooi and Renne (2011a)) can also formalize reasoning about information change produced by public and semi-private announcement. Different from other logical frameworks, AUL models information change by updating the epistemic access relation, without changing the domain of the model. This makes AUL very suitable for modeling information change on knowledge-how. Recall Example 5.1.1 in Chapter 5. If a new scientific discovery announces that the pill will certainly cure the pain no matter whether there is an allergy or not. In AUL, this announcement is formalized as removing the reflexive arrow on s_4. The doctor then will know that he can treat the patient by pill without any surgery. However, in AUL, it is common knowledge among agents how each will process incoming information. This assumption of common update policy is dropped in its extension, Generalized Arrow Update Logic (see Kooi and Renne (2011b)) (GAUL), which can capture the same information change that can be modeled in DEL.

Although DEL and GAUL are much more expressive than AUL, the great expressive power does not come for free. Their update operators are much more complex than the update operator of AUL. This chapter presents a variation of AUL, Private Arrow Update Logic (PAUL), which also drops the common update-policy assumption of AUL (so that private announcement can be expressed) and keeps the update operator as simple and intuitive as AUL. This logic framework is also inspired by the context-indexed semantics developed in Wang (2011) and Wang and Cao (2013). As we will see, PAUL can formalize reasoning about information change due to public, private and semi-private announcement.

The rest of the chapter is organized as follows: Section 6.2 proposes the language and semantics of PAUL, and works out some examples; Section 6.3 presents the tableau calculus for PAUL and show soundness and completeness; Section 6.4 shows PAUL is decidable; Section 6.5 concludes with some directions for further research.

6.2 The logic PAUL

6.2.1 Syntax and semantics

In this section, we introduce the language of this logic. This language differs from the language of AUL in the sense that each update information is only visible to an agent group.

Definition 6.2.1 (PAUL Language) *Let* $\mathbf{Agt}$ *be a nonempty finite set of agents, and let* $\mathbf{P}$ *be a countable set of atomic propositions. The PAUL language is generated by the following BNF:*

$$\varphi ::= \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid [U, G]\varphi \mid \Box_a \varphi$$
$$U ::= \{(\varphi, a, \varphi)\} \mid \{(\varphi, a, \varphi)\} \cup U$$

where $p \in \mathbf{P}$, $a \in \mathbf{Agt}$ *and* $G \subseteq \mathbf{Agt}$ *is a superset of the set of agents occurring in* U.

We will often omit parentheses around expressions when doing so ought not cause confusion. The expression φ is called a PAUL-formula (or just formula). The expression

$[U, G]$ occurring in a formula is called a PAUL-update (or just update), which consists of an update core U and an agent group G to which the update is visible. We let $\mathcal{L}_{\text{PAUL}}$ denote the set of formulas and updates. Given formulas φ and ψ and an agent $a \in \mathbf{Agt}$, the syntactic object $(\varphi, a, \psi) \in U$ is called an a-arrow specification. As usual, we use the following abbreviations: $\bot := \neg\top$, $\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)$, $\varphi \to \psi := \neg\varphi \vee \psi$, $\Diamond_a\varphi := \neg\Box_a\neg\varphi$.

Intuitively, the formula $\Box_a\varphi$ expresses that agent a believes φ. The formula $[U, G]\varphi$ expresses that φ holds after the arrow update $[U, G]$. The update $[U, G]$ means that the update is visible only to agents in G. Please note that the update in AUL has only one part, that is $[U]$, which is visible for all agents. Therefore, the update $[U]$ in AUL is the same as the update $[U, \mathbf{Agt}]$ here.

Definition 6.2.2 (Kripke Model) *A Kripke model $\mathcal{M}$ is a tuple $\langle W^{\mathcal{M}}, R^{\mathcal{M}}, V^{\mathcal{M}}\rangle$, consisting of a nonempty set $W^{\mathcal{M}}$ of worlds, a function $R^{\mathcal{M}}$ assigning each agent $a \in \mathbf{Agt}$ a binary relation $R_a^{\mathcal{M}} \subseteq W^{\mathcal{M}} \times W^{\mathcal{M}}$ ($R_a^{\mathcal{M}}$ can also be seen as a function from $W^{\mathcal{M}}$ to $2^{W^{\mathcal{M}}}$), and a function $V^{\mathcal{M}} : \mathbf{P} \to \mathcal{P}(W^{\mathcal{M}})$. A pointed Kripke model is a pair $(\mathcal{M}, s)$ consisting of a Kripke model $\mathcal{M}$ and a world $s \in W^{\mathcal{M}}$; the world s is called the point of $(\mathcal{M}, s)$.*

Given a Kripke model $\mathcal{M}$, we call $W^{\mathcal{M}}$ the domain of the model. For each agent $a \in \mathbf{Agt}$, we call $R_a^{\mathcal{M}}$ a's *possibility relation* since it defines what worlds agent a considers possible in any given world. Please note that updates considered in this chapter do not change any bare facts but only the agent's beliefs. Therefore, when an update happens, we do not have to change the domain of the model but only change the possibility relations (or 'arrows').

Definition 6.2.3 *Let $\rho = [U_1, G_1] \cdots [U_n, G_n]$ be an update sequence (or just sequence), and $\rho = \epsilon$ if $n = 0$. The update sequence $\rho|_a$ is defined by the following induction on n.*

$$\epsilon|_a = \epsilon$$

$$(\rho[U, G])|_a = \begin{cases} \rho|_a & a \notin G \\ \rho|_a[U, G] & a \in G \end{cases}$$

The sequence $\rho|_a$ means the updates visible to the agent a.

Definition 6.2.4 (PAUL Semantics) *Given a pointed Kripke model $(\mathcal{M}, s)$, an update sequence ρ and a formula φ, we write $\mathcal{M}, s \vDash_\rho \varphi$ to mean that φ is true at $\mathcal{M}, s$ after updates ρ, and we write $\mathcal{M}, s \nvDash_\rho \varphi$ for the negation of $\mathcal{M}, s \vDash_\rho \varphi$. The relation (notation: $\vDash_\rho$) is defined by the following induction on formula construction.*

$$\begin{array}{lll}
\mathcal{M}, s \vDash_\rho \top & & \\
\mathcal{M}, s \vDash_\rho p & \textit{iff} & s \in V(p) \\
\mathcal{M}, s \vDash_\rho \neg\varphi & \textit{iff} & \mathcal{M}, s \nvDash_\rho \varphi \\
\mathcal{M}, s \vDash_\rho (\varphi \wedge \psi) & \textit{iff} & \mathcal{M}, s \vDash_\rho \varphi \textit{ and } \mathcal{M}, s \vDash_\rho \psi \\
\mathcal{M}, s \vDash_\rho [U, G]\varphi & \textit{iff} & \mathcal{M}, s \vDash_{\rho[U,G]} \varphi \\
\mathcal{M}, s \vDash_\rho \Box_a\varphi & \textit{iff} & \forall t \in W^{\mathcal{M}} : (s, t) \in R_a^{\mathcal{M}} * (\rho|_a) \textit{ implies } \mathcal{M}, t \vDash_{\rho|_a} \varphi \\
R_a^{\mathcal{M}} * \epsilon & \stackrel{\textit{def}}{=} & R_a^{\mathcal{M}} \\
R_a^{\mathcal{M}} * (\rho'[U, G]) & \stackrel{\textit{def}}{=} & \{(s, t) \in R_a^{\mathcal{M}} * \rho' \mid \textit{ there exists } (\varphi, a, \psi) \in U : \\
& & \mathcal{M}, s \vDash_{\rho'} \varphi \textit{ and } \mathcal{M}, t \vDash_{\rho'} \psi\}
\end{array}$$

We also write $\mathcal{M}, s \vDash_\epsilon \varphi$ as $\mathcal{M}, s \vDash \varphi$. To say that a formula φ is valid, written as $\vDash \varphi$, means that $\mathcal{M}, s \vDash \varphi$ for each pointed Kripke model $(\mathcal{M}, s)$. The negation of $\vDash \varphi$ is written as $\nvDash \varphi$. To say that a formula φ is satisfiable means there exists a pointed model $(\mathcal{M}, s)$ such that $\mathcal{M}, s \vDash \varphi$.

The binary relation $R_a^{\mathcal{M}} * \rho|_a$ is a's possibility relation after the announcement sequence ρ. Compared to product semantics, such as in DEL and GAUL, the context-indexed semantics here has the following characteristics. Firstly, we know that updates change only agents' beliefs but not bare facts. This feature is more clear in this semantics because only the possibility relation is updated when $\Box$-formulas are evaluated. Moreover, product semantics always update the domain of the model when an update happens, which means that the size of the model may grow rapidly along with the length of update sequence ρ, but this is not the case here. This is because when $\Box_a$-formulas are evaluated, we do not change the domain of the model but only update a's possibility relation with respect to the update sequence visible to a, namely $\rho|_a$.

Kooi and Renne (2011a) present an axiomatic theory for AUL, in which the most important axiom states that an agent's belief after an update can be reduced to his (her) belief before the update. The following proposition shows that the PAUL version of this reduction axiom also holds.

Proposition 6.2.5 *$\vDash [U, G]\Box_a\varphi \leftrightarrow \bigwedge_{(\psi,a,\chi)\in U}(\psi \to \Box_a(\chi \to [U, G]\varphi))$ if $a \in G$.*

PROOF Let $(\mathcal{M}, s)$ be a pointed Kripke model. Firstly, we show that if $\mathcal{M}, s \vDash [U, G]\Box_a\varphi$ then $\mathcal{M}, s \vDash \bigwedge_{(\psi,a,\chi)\in U}(\psi \to \Box_a(\chi \to [U, G]\varphi))$. In order to show that $\mathcal{M}, s \vDash \bigwedge_{(\psi,a,\chi)\in U}(\psi \to \Box_a(\chi \to [U, G]\varphi))$, we only need to show that $\mathcal{M}, t \vDash [U, G]\varphi$ if there are $(\psi, a, \chi) \in U$ and $t \in R_a^{\mathcal{M}}(s)$ such that $\mathcal{M}, s \vDash \psi$ and $\mathcal{M}, t \vDash \chi$. Following by the assumption of $\mathcal{M}, s \vDash [U, G]\Box_a\varphi$, we then have $\mathcal{M}, t \vDash_{[U,G]} \varphi$. Therefore, we have $\mathcal{M}, t \vDash [U, G]\varphi$.

Secondly, we show that if $\mathcal{M}, s \vDash \bigwedge_{(\psi,a,\chi)\in U}(\psi \to \Box_a(\chi \to [U, G]\varphi))$ then $\mathcal{M}, s \vDash [U, G]\Box_a\varphi$. Assume that $\mathcal{M}, s \nvDash [U, G]\Box_a\varphi$. It follows that there exists $t \in W^{\mathcal{M}}$ such that $(s, t) \in R_a^{\mathcal{M}} * [U, G]$ and $\mathcal{M}, t \nvDash_{[U,G]} \varphi$. Since $(s, t) \in R_a^{\mathcal{M}} * [U, G]$, it follows that $(s, t) \in R_a^{\mathcal{M}}$ and there exists $(\psi, a, \chi) \in U$ such that $\mathcal{M}, s \vDash \psi$ and $\mathcal{M}, t \vDash \chi$. Moreover, since $\mathcal{M}, s \vDash \bigwedge_{(\psi,a,\chi)\in U}(\psi \to \Box_a(\chi \to [U, G]\varphi))$, we then have $\mathcal{M}, t \vDash [U, G]\varphi$, namely $\mathcal{M}, t \vDash_{[U,G]} \varphi$. This is in contradiction with $\mathcal{M}, t \nvDash_{[U,G]} \varphi$. Therefore, we have if $\mathcal{M}, s \vDash \bigwedge_{(\psi,a,\chi)\in U}(\psi \to \Box_a(\chi \to [U, G]\varphi))$ then $\mathcal{M}, s \vDash [U, G]\Box_a\varphi$. $\Box$

The following proposition shows that if an update is not visible for an agent, then her belief after the update is the same as her belief before the update.

Proposition 6.2.6 *$\vDash [U, G]\Box_a\varphi \leftrightarrow \Box_a\varphi$ if $a \notin G$.*

PROOF We have the following:

$$
\begin{aligned}
&\mathcal{M}, s \vDash [U, G]\Box_a\varphi \\
\Leftrightarrow &\mathcal{M}, s \vDash_{[U,G]} \Box_a\varphi \\
\Leftrightarrow &\text{ for all } (s, t) \in R_a * ([U, G]|_a) : \mathcal{M}, t \vDash_{[U,G]|_a} \varphi \\
\Leftrightarrow &\text{ for all } (s, t) \in R_a : \mathcal{M}, t \vDash \varphi \qquad \text{due to } [U, G]|_a = \epsilon
\end{aligned}
$$

$$\Leftrightarrow \mathcal{M}, s \vDash \Box_a \varphi$$

$\square$

6.2.2 Announcements in PAUL

In this section, we will show how public, private and semi-private announcement are captured in PAUL. Let us consider the following scenario of a concealed coin, which is a tweaked version of an example used in Baltag and Moss (2004).

Example 6.2.7 (Basic scenario) *Two agents a and b enter a large room which contains a remote-controlled mechanical coin flipper. One of them presses a button, and the coin spins through the air and lands in a small box on a table with heads or tails lying up. The box is closed and they are too far away to see the coin.*

$$\begin{tikzcd} \underset{p}{H} \arrow[loop left, "{a,b}"] \arrow[rr, leftrightarrow, "{a,b}"] & & \underset{\neg p}{T} \arrow[loop right, "{a,b}"] \end{tikzcd}$$

Figure 6.1: the basic model $\mathcal{M}$

Just as in Baltag and Moss (2004), this can be modelled by a Kripke model $\mathcal{M}$, which is pictured in Figure 6.1. The possible world $H \in W^{\mathcal{M}}$ represents the possible fact that the coin is lying heads up, and $T \in W^{\mathcal{M}}$ represents tails up. The proposition p means that the coin is lying heads up, so it is only true in H. The possibility relations of a and b indicate that both of them do not know whether the coin is lying heads or tails up.

Example 6.2.8 (Public announcement) *After the basic scenario, one of them opens the box and puts the coin on the table for both to see. The effect of this event on their beliefs is the same as that of a* truthful *statement publicly announced to them that the coin is lying heads or tails up.*

$$\begin{tikzcd} \underset{p}{H} \arrow[loop left, "{a,b}"] & & \underset{\neg p}{T} \arrow[loop right, "{a,b}"] \end{tikzcd}$$

Figure 6.2: $R_a^{\mathcal{M}} * ([U_1, G_1]|_a)$ and $R_b^{\mathcal{M}} * ([U_1, G_1]|_b)$

After the truthful announcement that the coin is lying heads or tails up, both of them think there is only one possibility in any given world. Thus only their epistemic accesses to any given world should be preserved. This announcement is visible for both, since it is publicly announced. Therefore, this public and truthful announcement can be captured by the update $[U_1, G_1]$ where $U_1 = \{(p, a, p), (\neg p, a, \neg p), (p, b, p), (\neg p, b, \neg p)\}$ and $G_1 = \{a, b\}$. After the update $[U_1, G_1]$, the possibility relations of a and b turn out to be as shown in Figure 6.2. Moreover, since the update is visible to both of them, a's possibility relation in b's opinion is the same as a's real possibility relation, namely $R_a^{\mathcal{M}} * ([U_1, G_1]|_b|_a) = R_a^{\mathcal{M}} * ([U_1, G_1]|_a)$. If H is the actual world, after this public and truthful announcement, both of them believe that the coin is lying heads up and that the other also believes so. We can check the following formulas.

- $\mathcal{M}, H \vDash [U_1, G_1](\Box_a p \wedge \Box_b p)$
- $\mathcal{M}, H \vDash [U_1, G_1](\Box_a \Box_b p \wedge \Box_b \Box_a p)$

Example 6.2.9 (Private announcement) *After the basic scenario of Example 6.2.7, a secretly opens the box herself. Agent b does not observe that a opens the box, and indeed a is certain that b does not suspect that anything happened. The effect of this on their beliefs is the same as secretly and privately announcing the truth to a.*

$$\begin{tikzcd} H \arrow[loop left, "a"] & & T \arrow[loop right, "a"] \end{tikzcd}$$

(a) $R_a^{\mathcal{M}} * ([U_2, G_2]|_a)$

$$\begin{tikzcd} H \arrow[loop left, "b"] \arrow[r, leftrightarrow, "b"] & T \arrow[loop right, "b"] \end{tikzcd}$$

(b) $R_b^{\mathcal{M}} * ([U_2, G_2]|_b)$

$$\begin{tikzcd} H \arrow[loop left, "a"] \arrow[r, leftrightarrow, "a"] & T \arrow[loop right, "a"] \end{tikzcd}$$

(c) $R_a^{\mathcal{M}} * ([U_2, G_2]|_b|_a)$

Figure 6.3: The possibility relations after the update $[U_2, G_2]$

After the truth is announced to a, she thinks that there is only one possibility from any given world. Thus a's epistemic accesses to the world itself should be preserved after the announcement. Since the announcement is secret and private, it is visible only to a. This private and truthful announcement can be captured by the update $[U_2, G_2]$ which is defined as $U_2 = \{(p, a, p), (\neg p, a, \neg p)\}$ and $G_2 = \{a\}$.

After the update $[U_2, G_2]$, a's possibility relation (Figure 6.3a) will change, but b's possibility relation (Figure 6.3b) will remain the same as before. Moreover, since b does not suspect that anything happened, a's possibility relation in b's opinion (Figure 6.3c) does not change at all after the announcement. After this private and truthful announcement to a, only a believes the truth while nothing happened to b's beliefs. We can check the following formulas.

- $\mathcal{M}, H \vDash [U_2, G_2](\Box_a p \wedge \neg \Box_b p)$
- $\mathcal{M}, H \vDash [U_2, G_2]\neg \Box_b(\Box_a p \vee \Box_a \neg p)$

Example 6.2.10 (Semi-private announcement) *After the basic scenario of Example 6.2.7, agent a opens the box herself. Agent b observes that a opens the box but does not see the coin. Agent a also does not disclose whether it is heads or tails. The effect of this on their beliefs is the same as a semi-private announcement to a, which means that the truth is announced to a only, but b notices what happened.*

$$\begin{tikzcd} H \arrow[loop left, "a"] & & T \arrow[loop right, "a"] \end{tikzcd}$$

(a) $R_a^{\mathcal{M}} * ([U_3, G_3]|_a)$

$$\begin{tikzcd} H \arrow[loop left, "b"] \arrow[r, leftrightarrow, "b"] & T \arrow[loop right, "b"] \end{tikzcd}$$

(b) $R_b^{\mathcal{M}} * ([U_3, G_3]|_b)$

$$\begin{tikzcd} H \arrow[loop left, "a"] & & T \arrow[loop right, "a"] \end{tikzcd}$$

(c) $R_a^{\mathcal{M}} * ([U_3, G_3]|_b|_a)$

Figure 6.4: The possibility relations after the update $[U_3, G_3]$

Since the truth is announced to a, she will know the truth after the announcement. The situation of b is a little complex. Firstly, b's possibility relation will remain the same as before since b is not announced the truth. Secondly, a's possibility relation in b's opinion will change since he observed that a is announced the truth. This semi-private announcement can be captured by the update $[U_3, G_3]$ which is defined as $U_3 = \{(p, a, p), (\neg p, a, \neg p), (\top, b, \top)\}$ and $G_3 = \{a, b\}$.

Agent a's possibility relation (Figure 6.4a) will change to the reflexive relation after the update. Since the announcement is not disclosed to b, b's possibility relation (Figure 6.4b) will not change after the update. However, after the update, a's possibility relation in b's opinion (Figure 6.4c) will change because b observes the announcement. After the announcement, b believes that a believes the truth, but b still could not distinguish between the fact that a believes p and the fact that a believes $\neg p$. We can check the following formulas.

- $\mathcal{M}, H \vDash [U_3, G_3](\Box_a p \wedge \neg\Box_b p)$
- $\mathcal{M}, H \vDash [U_3, G_3]\Box_b(\Box_a p \vee \Box_a \neg p)$

6.3 Tableau method

This section will present a proof method for PAUL that uses analytic tableaux. As a typical tableau method, given a formula φ, it systematically tries to construct a model for it. When it fails, φ is inconsistent and thus its negation is valid.

The tableau method in this chapter will manipulate tableau terms, which consist of two parts: the first part is an update sequence; the second part is a formula, or a check mark, or a cross mark. In addition, each term is prefixed by a label which stands for a possible world in the model under construction. A similar method is used in Fitting (1983); Massacci (2000); Balbiani et al. (2010); Aucher and Schwarzentruber (2013).

Definition 6.3.1 (Term) *A term (or tableau term) is a pair (ρ, x) where ρ is a finite update sequence $[U_1, G_1] \cdots [U_n, G_n]$ ($\rho = \epsilon$ if $n = 0$) and x is a check mark ✓, a cross mark ✗ or a formula $\varphi \in \mathcal{L}_{PAUL}$.*

Definition 6.3.2 (Length of term) *The length of a formula is defined as follows:*

$$l(p) = 1$$
$$l(\neg\varphi) = l(\varphi) + 1$$

$$
\begin{aligned}
l(\varphi \wedge \psi) &= l(\varphi) + l(\psi) + 1 \\
l(\Box_a \varphi) &= l(\varphi) + 1 \\
l([U, G]\varphi) &= l(U) + |G| + l(\varphi) + 1 \\
l(U) &= \sum\nolimits_{(\psi,a,\chi)\in U} (l(\psi) + l(\chi))
\end{aligned}
$$

The length of an update sequence is defined as follows:

$$l(\epsilon) = 0;\ l(\rho[U, G]) = l(\rho) + l(U) + |G|.$$

The length of a term is defined as follows:

$$l(\rho, \text{✗}) = l(\rho, \checkmark) = l(\rho);\ l(\rho, \varphi) = l(\rho) + l(\varphi).$$

Please note that the length of the term (ϵ, φ) is the same as the length of φ.

Definition 6.3.3 (Labelled term) *A* label *is an alternating sequence of integers and agents, namely $\sigma ::= n \mid \sigma a n$ where $n \in \mathbb{N}$ and $a \in$* **Agt***. A* labelled term *is a pair consisting of a label and a term, and we also write it as a triple $\langle \sigma, \rho, x \rangle$.*

Each label represents a possible world in a Kripke model. Moreover, a label $\sigma a n$ occurring on a branch of a tableau also indicates that there is an a-arrow from the possible world σ to the possible world $\sigma a n$. A labelled term $\langle \sigma, \rho, \varphi \rangle$ means φ is true at the possible world σ after the announcements ρ. A labelled term $\langle \sigma a n, \rho, \checkmark \rangle$ means the a-arrow from σ to $\sigma a n$ is preserved after the update sequence ρ. Conversely, a labelled term $\langle \sigma a n, \rho, \text{✗} \rangle$ means the a-arrow is not preserved.

Definition 6.3.4 (Branch) *A branch is a set of labelled terms. A label σ is new in a branch b if there is no term in b that is labelled with σ.*

Definition 6.3.5 (Tableau) *A tableau for $\varphi_0 \in \mathcal{L}$ is a set of branches inductively defined as follows.*

- $T = \{\{(0, \epsilon, \varphi_0)\}\}$*. This is called the initial tableau for φ_0.*
- $T = (T' \setminus \{b\}) \cup B$*, where T' is a tableau for φ_0 that contains the branch b and B is a finite set of branches generated by applying one of the tableau rules in Table 6.1 on b. For instance, let $b = \{\langle \sigma, \rho, \neg(\varphi \wedge \psi) \rangle\}$ then $B = \{b \cup \{\langle \sigma, \rho, \neg\varphi \rangle\}, b \cup \{\langle \sigma, \rho, \neg\psi \rangle\}\}$*

Rules $(\neg\neg)$, $(\neg\wedge)$ and $(\wedge)$ are exactly as for propositional logic. Rules $(\neg\Box_a)$ and $\Box_a$ are different from their counterparts commonly used in tableau calculi for normal modal logic. The intuition behind Rule $(\neg\Box_a)$ is that if the possible world that σ stands for satisfied $\neg\Box_a\varphi$ after the update sequence ρ then it needs to satisfy the following conditions: there exists a possible world that is represented by $\sigma a n$ (the form of $\sigma a n$ indicates that there is an a-arrow from σ to $\sigma a n$); $\langle \sigma a n, \rho|_a, \checkmark \rangle$ means the a-arrow from σ to $\sigma a n$ will be preserved after the update sequence $\rho|_a$; $\langle \sigma a n, \rho|_a, \neg\varphi \rangle$ means $\neg\varphi$ is true in $\sigma a n$ after the update sequence $\rho|_a$. Similarly, Rule $(\Box_a)$ means that $\Box_a\varphi$ is true in σ after ρ if and only if for each possible world that is accessible by a-arrow from σ: either the a-arrow is removed after $\rho|_a$, or φ is true in it after $\rho|_a$.

$$(\neg\neg)\ \frac{\langle \sigma, \rho, \neg\neg\varphi\rangle}{\langle \sigma, \rho, \varphi\rangle}$$

$$(\neg\wedge)\ \frac{\langle \sigma, \rho, \neg(\varphi \wedge \psi)\rangle}{\langle \sigma, \rho, \neg\varphi\rangle \mid \langle \sigma, \rho, \neg\psi\rangle} \qquad (\wedge)\ \frac{\langle \sigma, \rho, \varphi \wedge \psi\rangle}{\begin{array}{c}\langle \sigma, \rho, \varphi\rangle\\ \langle \sigma, \rho, \psi\rangle\end{array}}$$

$$(\neg\Box_a)\ \frac{\langle \sigma, \rho, \neg\Box_a\varphi\rangle}{\begin{array}{c}\langle \sigma an, \rho|_a, \checkmark\rangle\\ \langle \sigma an, \rho|_a, \neg\varphi\rangle\end{array}}\ \sigma an \text{ is new.}$$

$$(\Box_a)\ \frac{\langle \sigma, \rho, \Box_a\varphi\rangle}{\langle \sigma an, \rho|_a, \varphi\rangle \mid \langle \sigma an, \rho|_a, \mathsf{X}\rangle}\ \sigma an \text{ is used.}$$

$$(\neg[U,G])\ \frac{\langle \sigma, \rho, \neg[U,G]\varphi\rangle}{\langle \sigma, \rho[U,G], \neg\varphi\rangle} \qquad ([U,G])\ \frac{\langle \sigma, \rho, [U,G]\varphi\rangle}{\langle \sigma, \rho[U,G], \varphi\rangle}$$

$$(\checkmark_1)\ \frac{\langle \sigma an, \rho[U,G], \checkmark\rangle}{\left.\begin{array}{c}\langle \sigma an, \rho, \checkmark\rangle\\ \langle \sigma, \rho, \psi_1\rangle\\ \langle \sigma an, \rho, \chi_1\rangle\end{array}\right| \ldots \left|\begin{array}{c}\langle \sigma an, \rho, \checkmark\rangle\\ \langle \sigma, \rho, \psi_k\rangle\\ \langle \sigma an, \rho, \chi_k\rangle\end{array}\right.}\ (\psi_i, a, \chi_i) \in U \text{ for each } 1 \leq i \leq k$$

$$(\checkmark_2)\ \frac{\langle \sigma an, \rho[U,G], \checkmark\rangle}{\begin{array}{c}\langle \sigma an, \epsilon, \checkmark\rangle\\ \langle \sigma an, \epsilon, \mathsf{X}\rangle\end{array}}\ \text{there are no } \psi \text{ and } \chi \text{ such that } (\psi, a, \chi) \in U$$

$$(\mathsf{X}_1)\ \frac{\langle \sigma an, \rho[(\psi, a, \chi), G], \mathsf{X}\rangle}{\langle \sigma an, \rho, \mathsf{X}\rangle \mid \langle \sigma, \rho, \neg\psi\rangle \mid \langle \sigma an, \rho, \neg\chi\rangle}$$

$$(\mathsf{X}_2)\ \frac{\langle \sigma an, \rho[U,G], \mathsf{X}\rangle}{\begin{array}{c}\langle \sigma an, \rho[(\psi_1, a_1, \chi_1), G], \mathsf{X}\rangle\\ \vdots\\ \langle \sigma an, \rho[(\psi_k, a_k, \chi_k), G], \mathsf{X}\rangle\end{array}}\ \begin{array}{l}|U| = k,\ k \geq 2 \text{ and } (\psi_i, a_i, \chi_i) \in U\\ \text{for each } 1 \leq i \leq k\end{array}$$

Table 6.1: Tableau rules

1.	$\langle 0, \epsilon, [(q, b, q), b]\Box_a p \wedge \neg\Box_a p\rangle$	
2.	$\langle 0, \epsilon, [(q, b, q), b]\Box_a p\rangle$	(Rule $(\wedge)$: 1)
3.	$\langle 0, \epsilon, \neg\Box_a p\rangle$	(Rule $(\wedge)$: 1)
4.	$\langle 0, [(q, b, q), b], \Box_a p\rangle$	(Rule $([U, G])$: 2)
5.	$\langle 0a1, \epsilon, \checkmark\rangle$	(Rule $(\neg\Box_a)$: 3)
6.	$\langle 0a1, \epsilon, \neg p\rangle$	(Rule $(\neg\Box_a)$: 3)

7.	$\langle 0a1, \epsilon, p\rangle$	(Rule $(\Box_a)$: 4)	8.	$\langle 0a1, \epsilon, \mathsf{X}\rangle$	(Rule $(\Box_a)$: 4)
	closed	(6, 7)		closed	(5, 8)

Figure 6.5: Closed tableau for the formula $[(q, b, q), b]\Box_a p \wedge \neg\Box_a p$

Rule $(\neg[U, G])$ and Rule $([U, G])$ reflect the feature of the semantics that the updates are just remembered and they are used to update the possibility relation only when $\Box_a$ formulas are evaluated. Rule $(\checkmark_1)$ means that the a-arrow is preserved after $\rho[U, G]$ if and only if it is firstly preserved after ρ and then preserved by some a-arrow specification in U. Rule $(\checkmark_2)$ says it is not possible that the a-arrow from σ to σan is preserved after $\rho[U, G]$ if there is no a-arrow specifications in U. Rule (X_1) and Rule (X_2) specify the conditions under which the a-arrow from σ to σan will be removed. It is removed after $\rho[(\psi, a, \chi), G]$ if either it is already removed after ρ, or it cannot be preserved the specification (ψ, a, χ). Rule (X_2) corresponds to the semantics that $R_a * (\rho[U, G]) = \bigcup_{(\psi,a,\chi)\in U} R_a * (\rho[(\psi, a, \chi), G])$. Please note that it is trivially true that any a-arrow will be remove after $\rho[(\psi, a', \chi), G]$ if $a' \neq a$.

The following proposition is obvious according to the tableau rules.

Proposition 6.3.6 *Given a tableau T and a branch $b \in T$, if $\langle\sigma, \rho, x\rangle \in b$ and $x = \checkmark/\mathsf{X}$ then $\sigma = \sigma' an$ for some label σ', $a \in A$ and $n \in \mathbb{N}$.*

Definition 6.3.7 (Closed tableau) *A branch b is* closed *if and only if we have either $\{\langle\sigma, \rho, p\rangle, \langle\sigma, \rho', \neg p\rangle\} \subseteq b$ for some σ, ρ, ρ' and p, or $\{\langle\sigma an, \epsilon, \checkmark\rangle, \langle\sigma an, \epsilon, \mathsf{X}\rangle\} \subseteq b$ for some σan, otherwise it is* open. *A tableau is* closed *if and only if all its branches are closed, otherwise it is* open.

Example 6.3.8 *In Figure 6.5, the tableau method is used to show the validity of one instance of the formula of Proposition 6.2.6. The rightmost column shows which tableau rule is applied in each line.*

Next, we will show the soundness, but first we need another definition.

Definition 6.3.9 (Satisfiable branch) *Given a Kripke model $\mathcal{M}$ and a branch b, let f be a function from the labels used in b to $W^{\mathcal{M}}$. We say b is satisfied by $\mathcal{M}$ and f if the followings hold.*

- $\mathcal{M}, f(\sigma) \vDash_\rho \varphi$ *for each* $\langle\sigma, \rho, \varphi\rangle \in b$;
- $(f(\sigma), f(\sigma an)) \in R_a^{\mathcal{M}} * \rho$ *for each* $\langle\sigma an, \rho, \checkmark\rangle \in b$;
- $(f(\sigma), f(\sigma an)) \notin R_a^{\mathcal{M}} * \rho$ *for each* $\langle\sigma an, \rho, \mathsf{X}\rangle \in b$.

If there are such Kripke models and functions, we say b is satisfiable.

It is obvious that if b is satisfiable then all φ with $(\sigma, \epsilon, \varphi) \in b$ is satisfiable.

Theorem 6.3.10 (Soundness) *If there is a closed tableau for $\neg\varphi_0$, then φ_0 is valid.*

PROOF We show that if φ_0 is satisfiable then there is no closed tableau for φ_0. If φ_0 is satisfiable, it is obvious that the branch in the initial tableau for φ_0 is satisfiable. Therefore, it is enough to show that all tableau rules preserve satisfiability, that is, if a branch b is satisfiable then at least one branch in B is satisfiable where B is the branch set generated by applying a tableau rule on b.

Suppose that there are a Kripke model $\mathcal{M}$ and a function f by which the branch b is satisfied. Next we will show all tableau rules in Table 6.1 that are applicable to b preserve satisfiability. The rules $(\neg\neg)$, $(\neg\wedge)$ and $(\wedge)$ are analogues to the rules commonly used for propositional logic; we restrict our attention to the other rules.

1. Rule $(\neg\Box_a)$: By the assumption, we have that $\langle\sigma, \rho, \neg\Box_a\varphi\rangle \in b$, $B = \{b \cup \{\langle\sigma an, \rho|_a, \checkmark\rangle, \langle\sigma an, \rho|_a, \neg\varphi\rangle\}\}$ where σan is new in b. Since $\mathcal{M}, f(\sigma) \vDash_\rho \varphi$, it follows that there exists $t \in W^{\mathcal{M}}$ such that $(f(\sigma), t) \in R_a^{\mathcal{M}} * \rho|_a$ and $\mathcal{M}, t \vDash_{\rho|_a} \neg\varphi$. Now consider the function f' such that $f'(\sigma') = f(\sigma')$ for all σ' used in b and $f'(\sigma an) = t$. We then have $(f'(\sigma), f'(\sigma an)) \in R_a^{\mathcal{M}} * \rho|_a$ and $\mathcal{M}, f'(\sigma an) \vDash_{\rho|_a} \neg\varphi$. Therefore, the branch in B is satisfiable.

2. Rule $(\Box_a)$: For each σan which is used in b, we have either $(f(\sigma), f(\sigma an)) \in R_a^{\mathcal{M}} * \rho|_a$ or $(f(\sigma), f(\sigma an)) \notin R_a^{\mathcal{M}} * \rho|_a$. It follows by $\langle\sigma, \rho, \Box_a\varphi\rangle \in b$ that $\mathcal{M}, f(\sigma) \vDash_\rho \Box_a\varphi$. If $(f(\sigma), f(\sigma an)) \in R_a^{\mathcal{M}} * \rho|_a$ then we have $\mathcal{M}, f(\sigma an) \vDash_{\rho|_a} \varphi$. Thus, the branch $b \cup \{\langle\sigma an, \rho|_a, \varphi\rangle\}$ is satisfiable. If $(f(\sigma), f(\sigma an)) \notin R_a^{\mathcal{M}} * \rho|_a$, the branch $b \cup \{\langle\sigma an, \rho|_a, \mathsf{X}\rangle\}$ is satisfiable.

3. Rule $(\neg[U, G])$: Since $\langle\sigma, \rho, \neg[U, G]\varphi\rangle \in b$, it follows by the assumption that $\mathcal{M}, f(\sigma) \vDash_\rho \neg[U, G]\varphi$. By semantics, it follows that $\mathcal{M}, f(\sigma) \vDash_\rho \neg[U, G]\varphi$ iff $\mathcal{M}, f(\sigma) \nvDash_\rho [U, G]\varphi$ iff $\mathcal{M}, f(\sigma) \nvDash_{\rho[U,G]} \varphi$ iff $\mathcal{M}, f(\sigma) \vDash_{\rho[U,G]} \neg\varphi$. Since $B = \{b \cup \{\langle\sigma, \rho[U, G], \neg\varphi\rangle\}\}$, the branch in B is satisfiable.

4. Rule $([U, G])$: Since $\langle\sigma, \rho, [U, G]\varphi\rangle \in b$, it follows by the assumption that $\mathcal{M}, f(\sigma) \vDash_\rho [U, G]\varphi$. By semantics, it follows that $\mathcal{M}, f(\sigma) \vDash_\rho [U, G]\varphi$ iff $\mathcal{M}, f(\sigma) \vDash_{\rho[U,G]} \varphi$. Because of $B = \{b \cup \{\langle\sigma, \rho[U, G], \varphi\rangle\}\}$, thus the branch in B is satisfiable.

5. Rule $(\checkmark_1)$: Since we have $\langle\sigma an, \rho[U, G], \checkmark\rangle \in b$, it follows by the assumption that $(f(\sigma), f(\sigma an)) \in R_a^{\mathcal{M}} * \rho[U, G]$. It follows by semantics that $(f(\sigma), f(\sigma an)) \in R_a^{\mathcal{M}} * \rho$ and there exists $(\psi, a, \chi) \in U$ such that $\mathcal{M}, f(\sigma) \vDash_\rho \psi$ and $\mathcal{M}, f(\sigma an) \vDash_\rho \chi$. Let $b' = b \cup \{\langle\sigma an, \rho, \checkmark\rangle, \langle\sigma, \rho, \psi\rangle, \langle\sigma an, \rho, \chi\rangle\}$, we then have b' is satisfiable. Due to $b' \in B$, thus one branch in B is satisfiable.

6. Rule $(\checkmark_2)$: Since b is satisfied by $\mathcal{M}$ and f, it is obvious that Rule $(\checkmark_2)$ is not applicable to b.

7. Rule (X_1): Since $\langle\sigma an, \rho[(\psi, a, \chi), G], \mathsf{X}\rangle \in b$, it follows by the assumption that $(f(\sigma), f(\sigma an)) \notin R_a^{\mathcal{M}} * \rho[(\psi, a, \chi), G]$. It follows by the semantics that $(f(\sigma), f(\sigma an)) \notin R_a^{\mathcal{M}} * \rho$ or $\mathcal{M}, f(\sigma) \nvDash_\rho \psi$ or $\mathcal{M}, f(\sigma an) \nvDash_\rho \chi$. Therefore, we have that one branch of $B = \{b \cup \{\langle\sigma an, \rho, \mathsf{X}\rangle\}, b \cup \{\langle\sigma, \rho, \neg\psi\rangle\}, b \cup \{\langle\sigma an, \rho, \neg\chi\rangle\}\}$ is satisfiable.

8. Rule (X_2): Since $\langle \sigma an, \rho[U,G], \mathsf{X}\rangle \in b$, it follows that $(f(\sigma), f(\sigma an)) \notin R_a^{\mathcal{M}} * \rho[U,G]$. It follows by the semantics that $R_a^{\mathcal{M}} * \rho[U,G] = \bigcup_{(\psi,a,\chi)\in U} R_a^{\mathcal{M}} * \rho[(\psi,a,\chi),G]$. Therefore, we have $(f(\sigma), f(\sigma an)) \notin R_a^{\mathcal{M}} * \rho[(\psi,a',\chi),G]$ for each $(\psi,' a,\chi) \in U$ and $a' = a$. If $(\psi,a',\chi) \in U$ and $a' \neq a$, it follows that $R_a^{\mathcal{M}} * (\rho[(\psi,a',\chi),G]) = \emptyset$. Thus we have that the branch $b \cup \{\langle \sigma an, \rho[(\psi,a',\chi),G], \mathsf{X}\rangle \mid (\psi,a',\chi) \in U\}$ is satisfiable.

$\square$

In the rest of the section, we prove completeness. First, we need another auxiliary definition.

Definition 6.3.11 (Saturated tableau) *A branch b is saturated if and only if it is saturated under all tableau rules, as defined below:*

1. *b is saturated under Rule $(\neg\neg)$ if and only if $\langle \sigma, \rho, \neg\neg\varphi\rangle \in b$ implies $\langle \sigma, \rho, \varphi\rangle \in b$;*
2. *b is saturated under Rule $(\neg\wedge)$ if and only if $\langle \sigma, \rho, \neg(\varphi\wedge\psi)\rangle \in b$ implies $\langle \sigma, \rho, \neg\varphi\rangle \in b$ or $\langle \sigma, \rho, \neg\psi\rangle \in b$;*
3. *b is saturated under Rule $(\wedge)$ if and only if $\langle \sigma, \rho, (\varphi\wedge\psi)\rangle \in b$ implies $\langle \sigma, \rho, \varphi\rangle \in b$ and $\langle \sigma, \rho, \psi\rangle \in b$;*
4. *b is saturated under Rule $(\neg\Box_a)$ if and only if $\langle \sigma, \rho, \neg\Box_a\varphi\rangle \in b$ implies that $\{\langle \sigma an, \rho|_a, \checkmark\rangle, \langle \sigma an, \rho|_a, \neg\varphi\rangle\} \subset b$ for some $n \in \mathbb{N}$;*
5. *b is saturated under Rule $(\Box_a)$ if and only if $\langle \sigma, \rho, \Box_a\varphi\rangle \in b$ implies that for each σan occurring in b we have $\langle \sigma an, \rho|_a, \mathsf{X}\rangle \in b$ or $\langle \sigma an, \rho|_a, \varphi\rangle \in b$;*
6. *b is saturated under Rule $(\neg[U,G])$ if and only if $\langle \sigma, \rho, \neg[U,G]\varphi\rangle \in b$ implies $\langle \sigma, \rho[UG], \neg\varphi\rangle \in b$;*
7. *b is saturated under Rule $([U,G])$ if and only if $\langle \sigma, \rho, [U,G]\varphi\rangle \in b$ implies $\langle \sigma, \rho[UG], \varphi\rangle \in b$;*
8. *b is saturated under Rule $(\checkmark_1)$ if and only if $\langle \sigma an, \rho[U,G], \checkmark\rangle \in b$ implies $\{\langle \sigma an, \rho, \checkmark\rangle, \langle \sigma, \rho, \psi\rangle, \langle \sigma an, \rho, \chi\rangle\} \subset b$ for some $(\psi, a, \chi) \in U$;*
9. *b is saturated under Rule $(\checkmark_2)$ if and only if $\langle \sigma an, \rho[U,G], \checkmark\rangle \in b$ implies $\{\langle \sigma an, \epsilon, \checkmark\rangle, \langle \sigma, \epsilon, \mathsf{X}\rangle, \} \subset b$;*
10. *b is saturated under Rule (X_1) if and only if $\langle \sigma an, \rho[(\psi,a,\chi),G], \mathsf{X}\rangle \in b$ implies $\langle \sigma an, \rho, \mathsf{X}\rangle \in b$ or $\langle \sigma, \rho, \neg\psi\rangle \in b$ or $\langle \sigma an, \rho, \neg\chi\rangle \in b$.*
11. *b is saturated under Rule (X_2) if and only if $\langle \sigma an, \rho[U,G], \mathsf{X}\rangle \in b$ implies that $\{\langle \sigma an, \rho[(\psi,a',\chi),G], \mathsf{X}\rangle \mid (\psi,a',\chi) \in U\} \subset b$, where there are at least two specifications in U.*

We say a tableau is saturated *if and only if all its branches are saturated.*

The following two propositions are obvious by the tableau rules.

Proposition 6.3.12 *Given a saturated tableau T and a branch $b \in T$, if $\langle \sigma an, \rho, \checkmark \rangle \in b$ then $\langle \sigma an, \epsilon, \checkmark \rangle \in b$.*

Proposition 6.3.13 *Given a saturated tableau T and a branch $b \in T$, if a label σan occurs in b then $\langle \sigma an, \epsilon, \checkmark \rangle \in b$.*

Now, we are ready to prove the completeness.

Theorem 6.3.14 (completeness) *If φ_0 is valid, there is a closed tableau for $\neg\varphi_0$.*

PROOF We only need to show that if all tableaux for φ_0 are open then φ_0 is satisfiable. Since each tableau for φ_0 can be extended to be saturated and there is at least one tableau for φ_0, i.e. the initial tableau, there exists an open and saturated tableau for φ_0 if all its tableaux are open.

Let T be an open and saturated tableau for φ_0 and b be an open and saturated branch of T. In order to show φ_0 is satisfiable, we only need to show that the branch b is satisfiable in the sense of Definition 6.3.9. Next we will construct a model $\mathcal{M}^c$ and we will show that b is satisfied by $\mathcal{M}^c$. The model $\mathcal{M}^c = \langle W, R, V\rangle$ is defined as follows.

$$\begin{aligned} W &= \{\sigma \mid \sigma \text{ is used in } b\} \\ R_a &= \{(\sigma, \sigma an) \mid (\sigma an, \epsilon, \checkmark) \in b\} \text{ for each } a \in \mathbf{Agt} \\ V(p) &= \{\sigma \mid (\sigma, \rho, p) \in b \text{ for some } \rho\} \end{aligned}$$

Please note that if σan is used in b then so is σ.

By induction on the length of terms, we will show that b is satisfied by $\mathcal{M}^c$ and I, where I is the function $I(\sigma) = \sigma$. For abbreviation, we will write $I(\sigma)$ as σ. For the case of $l(\rho, x) = 0$, the term (ρ, x) can only be of the form (ϵ, X) or $(\epsilon, \checkmark)$. Furthermore, it cannot be of the form (ϵ, X). Assuming $(\sigma an, \epsilon, \mathsf{X}) \in b$ for some label σan, it follows by Proposition 6.3.13 that $(\sigma an, \epsilon, \checkmark) \in b$, this is in contradiction with that b is open. Therefore, in this case, we only need to show that $(\sigma, \sigma an) \in R_a$ for each σan with $(\sigma an, \epsilon, \checkmark) \in b$, which is obvious by the definition of the model $\mathcal{M}^c$.

With the inductive hypothesis that each labelled term $(\sigma, \rho, x) \in b$ with $l(\rho, x) < n$ satisfies the conditions declared in Definition 6.3.9, we will show that each labelled term $(\sigma, \rho, x) \in b$ with $l(\rho, x) = n$ also satisfies the conditions, where $n \geq 1$.

If x is a formula, there are different cases according to the form of the formula, as below:

1. $(\sigma, \rho, p) \in b$: It is obvious that $\mathcal{M}^c, \sigma \vDash_\rho p$.

2. $(\sigma, \rho, \neg p) \in b$: Assuming $\sigma \in V(p)$, it follows that $(\sigma, \rho', p) \in b$ for some ρ'. This is in contradiction with the assumption that b is open. Therefore, we have $\sigma \notin V(p)$, namely $\mathcal{M}^c, \sigma \vDash_\rho \neg p$.

3. $(\sigma, \rho, \neg\neg\varphi) \in b$: Since b is saturated, it follows that $(\sigma, \rho, \varphi) \in b$. Since $l(\rho, \varphi) < l(\rho, \neg\neg\varphi)$, it follows by IH that $\mathcal{M}^c, \sigma \vDash_\rho \varphi$. Therefore, we have $\mathcal{M}^c, \sigma \vDash_\rho \neg\neg\varphi$.

4. $(\sigma, \rho, \varphi \wedge \psi) \in b$: Since b is saturated, it follows that $(\sigma, \rho, \varphi) \in b$ and $(\sigma, \rho, \psi) \in b$. Since $l(\rho, \varphi), l(\rho, \psi) < l(\rho, \varphi \wedge \psi)$, it follows by IH that $\mathcal{M}^c, \sigma \vDash_\rho \varphi$ and $\mathcal{M}^c, \sigma \vDash_\rho \psi$. Therefore, we have $\mathcal{M}^c, \sigma \vDash_\rho \varphi \wedge \psi$.

5. $(\sigma, \rho, \neg(\varphi \wedge \psi)) \in b$: Since b is saturated, it follows that $(\sigma, \rho, \neg\varphi) \in b$ or $(\sigma, \rho, \neg\psi) \in b$. Since $l(\rho, \neg\varphi), l(\rho, \neg\psi) < l(\rho, \neg(\varphi \wedge \psi))$, it follows by IH that $\mathcal{M}^c, \sigma \vDash_\rho \neg\varphi$ or $\mathcal{M}^c, \sigma \vDash_\rho \neg\psi$. Therefore, we have $\mathcal{M}^c, \sigma \vDash_\rho \neg(\varphi \wedge \psi)$.

6. $(\sigma, \rho, \neg\Box_a\varphi) \in b$: Since b is saturated, it follows that $(\sigma an, \rho|_a, \neg\varphi) \in b$ and $(\sigma an, \rho|_a, \checkmark) \in b$ for some $n \in \mathbb{N}$. Since $l(\rho|_a, \neg\varphi), l(\rho|_a, \checkmark) < l(\rho, \neg\Box_a\varphi)$, it follows by IH that $(\sigma, \sigma an) \in R_a * (\rho|_a)$ and $\mathcal{M}^c, \sigma \vDash_{\rho|_a} \neg\varphi$. Therefore, we have $\mathcal{M}^c, \sigma \vDash_\rho \neg\Box_a\varphi$.

7. $(\sigma, \rho, \Box_a\varphi) \in b$: Let $\sigma' \in W$ be a state with $(\sigma, \sigma') \in R_a * (\rho|_a)$. In order to show $\mathcal{M}^c, \sigma \vDash_\rho \Box_a\varphi$, we need to show that $\mathcal{M}^c, \sigma' \vDash_{\rho|_a} \varphi$. Since $R_a * (\rho|_a) \subseteq R_a$, it follows that $\sigma' = \sigma an$ for some $n \in \mathbb{N}$. Assuming $(\sigma an, \rho|_a, \mathsf{X}) \in b$, it follows by IH that $(\sigma, \sigma an) \notin R_a * (\rho|_a)$. This is in contradiction with the assumption that $(\sigma, \sigma') \in R_a * (\rho|_a)$. Therefore, we have $(\sigma an, \rho|_a, \mathsf{X}) \notin b$. Since b is saturated, it follows that $(\sigma an, \rho|_a, \varphi) \in b$. It follows by IH that $\mathcal{M}^c, \sigma an \vDash_{\rho|_a} \varphi$.

8. $(\sigma, \rho, \neg[U, G]\varphi) \in b$: Since b is saturated, it follows that $(\sigma, \rho[U, G], \neg\varphi) \in b$. Since $l(\rho[U, G], \neg\varphi) < l(\rho, \neg[U, G]\varphi)$, it follows by IH that $\mathcal{M}^c, \sigma \vDash_{\rho[U,G]} \neg\varphi$. Therefore, we have $\mathcal{M}^c, \sigma \vDash_\rho \neg[U, G]\varphi$.

9. $(\sigma, \rho, [U, G]\varphi) \in b$: Since b is saturated, it follows that $(\sigma, \rho[U, G], \varphi) \in b$. Since $l(\rho[U, G], \varphi) < l(\rho, [U, G]\varphi)$, it follows by IH that $\mathcal{M}^c, \sigma \vDash_{\rho[U,G]} \varphi$. Therefore, we have $\mathcal{M}^c, \sigma \vDash_\rho [U, G]\varphi$.

If x in the term (ρ, x) is of the form $\checkmark$ or X, we have ρ is not ϵ because $l(\rho, x) \geq 1$. There are different cases, as below:

1. $(\sigma an, \rho[U, G], \checkmark) \in b$ and there exists an a-arrow specification in U: Since b is saturated, it follows that $\{\langle \sigma an, \rho, \checkmark \rangle, \langle \sigma, \rho, \psi \rangle, \langle \sigma an, \rho, \chi \rangle\} \subset b$ for some $(\psi, a, \chi) \in U$. Since $l(\rho, \checkmark), l(\rho, \psi), l(\rho, \chi) < l(\rho[U, G], \checkmark)$, it follows by IH that $(\sigma, \sigma an) \in R_a * \rho$, $\mathcal{M}^c, \sigma \vDash_\rho \psi$ and $\mathcal{M}^c, \sigma an \vDash_\rho \chi$. It follows that $(\sigma, \sigma an) \in R_a * (\rho[U, G])$.

2. $(\sigma an, \rho[U, G], \checkmark) \in b$ and there are no a-arrow specifications in U: Due to Rule $(\checkmark_2)$ and the fact that b is open and saturated, this case is impossible.

3. $(\sigma an, \rho[(\psi, a', \chi), G], \mathsf{X}) \in b$: If $a' \neq a$, it follows that $R_a * (\rho[(\psi, a', \chi), G]) = \emptyset$. It is obvious $(\sigma, \sigma an) \notin R_a * (\rho[(\psi, a', \chi), G])$. If $a' = a$, it follows by Rule (X_1) that $\langle \sigma an, \rho, \mathsf{X} \rangle \in b$, or $\langle \sigma, \rho, \neg\psi \rangle \in b$, or $\langle \sigma an, \rho, \neg\chi \rangle \in b$. Since $l(\rho, \mathsf{X}), l(\rho, \neg\psi), l(\rho, \neg\chi) < l(\rho[(\psi, a, \chi), G], \mathsf{X})$, it follows by IH that $(\sigma, \sigma an) \notin R_a * \rho$, or $\mathcal{M}^c, \sigma \vDash_\rho \neg\psi$, or $\mathcal{M}^c, \sigma an \vDash_\rho \neg\chi$. Each of them can derive that $(\sigma, \sigma an) \notin R_a * (\rho[(\psi, a, \chi), G])$.

4. $(\sigma an, \rho[U, G], \mathsf{X}) \in b$ and $|U| \geq 2$: If there are no a-arrow specifications in U, it is obvious that $(\sigma, \sigma an) \notin R_a * (\rho[U, G])$ since $R_a * (\rho[U, G]) = \emptyset$. Otherwise, let $(\psi_1, a, \chi_1), \cdots, (\psi_k, a, \chi_k)$ be all the a-arrow specifications in U. Since b is saturated, it follows by Rule (X_2) that $\langle \sigma an, \rho[(\psi_i, a, \chi_i), G], \mathsf{X} \rangle \in b$ for all $1 \leq i \leq k$. Since $l(\rho[(\psi_i, a, \chi_i), G], \mathsf{X}) < l([U, G], \mathsf{X})$ for all $1 \leq i \leq k$ due to $|U| \geq 2$, it follows by IH that $(\sigma, \sigma an) \notin R_a * (\rho[(\psi_i, a, \chi_i), G])$ for all $1 \leq i \leq k$. Since $R_a * (\rho[U, G]) = \bigcup_{1 \leq i \leq k} R_a * (\rho[(\psi_i, a, \chi_i), G])$, we have $(\sigma, \sigma an) \notin R_a * (\rho[U, G])$.

We have shown that all labelled terms in b satisfy the conditions declared in Definition 6.3.9. Since $\langle 0, \epsilon, \varphi_0 \rangle \in b$, thus we have $\mathcal{M}^c, 0 \models \varphi_0$. $\Box$

6.4 Decidability

In this section, we will show that PAUL is decidable, that is, the problem whether an PAUL formula φ is satisfiable can be answered in a finite number of steps. Our method is to show that PAUL has small model property. We will show that each satisfiable PAUL formula φ has a bounded small model in which φ is true. From the proof of Theorem 6.3.14, we have seen that we can construct a model for φ based on a saturated open branch if φ is satisfiable, and each state in the model is exactly a label used in the branch. Therefore, the key is to show that there are only finitely many labels used in the tableau branch.

For the commonly used tableau calculus for normal modal logic, each formula occurring in the tableau is a subformula of the destination formula, and this feature plays an important role to show the decidability of normal modal logic through tableau method. Similarly, we will define the notation of subterm here, and we will show that all terms occurring in the tableau are subterms.

Definition 6.4.1 (Subterm) *Given a term (ρ, x), the set of subterm of (ρ, x), denoted as $sub(\rho, x)$, is defined as below.*

$$\begin{aligned}
sub(\epsilon, ✗/✓) &= \{(\epsilon, ✗/✓)\} \\
sub(\rho[(\psi, a, \chi), G], ✗/✓) &= \{(\rho[(\psi, a, \chi), G], ✗/✓)\} \cup sub(\rho, \psi) \cup sub(\rho, \chi) \\
sub(\rho[U, G], ✗/✓) &= \{(\rho[U, G], ✗/✓)\} \cup \bigcup_{(\psi,a,\chi)\in U} sub(\rho[(\psi, a, \chi), G], ✗/✓) \\
&\qquad \textit{where } |U| \geq 2 \\
sub(\rho, p) &= \{(\rho, p)\} \cup sub(\rho, ✗) \cup sub(\rho, ✓) \\
sub(\rho, \neg\varphi) &= \{(\rho, \neg\varphi)\} \cup sub(\rho, \varphi) \\
sub(\rho, \varphi \wedge \psi) &= \{(\rho, \varphi \wedge \psi)\} \cup sub(\rho, \varphi) \cup sub(\rho, \psi) \\
sub(\rho, \Box_a\varphi) &= \{(\rho, \Box_a\varphi)\} \cup sub(\rho|_a, \varphi) \\
sub(\rho, [U, G]\varphi) &= \{(\rho, [U, G]\varphi)\} \cup sub(\rho[U, G], \varphi)
\end{aligned}$$

Let $sub^+(\rho, x)$ be the set $\{(\rho, \neg\varphi) \mid (\rho, \varphi) \in sub(\rho, x)\} \cup sub(\rho, x)$.

The following proposition states some properties of the subterm set.

Proposition 6.4.2 *We have the following results.*

- *$sub(\rho, x)$ is finite;*
- *$(\rho, ✗/✓) \in sub(\rho, \varphi)$;*
- *$(\rho, x) \in sub(\rho', x')$ implies $sub(\rho, x) \subseteq sub(\rho', x')$.*

Proposition 6.4.3 *Let T be a tableau for φ_0 and b be a branch of T. If $(\sigma, \rho, x) \in b$ then $(\rho, x) \in sub^+(\epsilon, \varphi_0)$.*

PROOF According to Definition 6.3.5, we prove this by induction on the process of construction of T. For the initial tableau $\{\{(0, \epsilon, \varphi_0)\}\}$, it is obvious. Next, we only need to show that all the tableau rules in Table 6.1 preserve the subterm property. The cases of the rules $(\neg\neg)$, $(\neg\wedge)$, $(\wedge)$, $([U, G])$ and (X_2) are obvious; we will restrict our attention to the other rules.

1. Rule $(\neg\Box_a)$: If $(\rho, \neg\Box_a\varphi) \in sub^+(\epsilon, \varphi_0)$, then we have $(\rho, \Box_a\varphi) \in sub(\epsilon, \varphi_0)$. Because $(\rho|_a, \varphi) \in sub(\rho, \Box_a\varphi)$, it follows by Proposition 6.4.2 that $(\rho|_a, \checkmark)$, $(\rho|_a, \varphi) \in sub^+(\epsilon, \varphi_0)$.

2. Rule $(\Box_a)$: If $(\rho, \Box_a\varphi) \in sub^+(\epsilon, \varphi_0)$, then we have $(\rho, \Box_a\varphi) \in sub(\epsilon, \varphi_0)$. Because $(\rho|_a, \varphi) \in sub(\rho, \Box_a\varphi)$, it follows by Proposition 6.4.2 that $(\rho|_a, \mathsf{X})$, $(\rho|_a, \varphi) \in sub^+(\epsilon, \varphi_0)$.

3. Rule $(\neg[U, G])$: If $(\rho, \neg[U, G]\varphi) \in sub^+(\epsilon, \varphi_0)$, then we have $(\rho, [U, G]\varphi) \in sub(\epsilon, \varphi_0)$. Since $(\rho[U, G], \varphi) \in sub(\rho, [U, G]\varphi)$, it follows by Proposition 6.4.2 that $(\rho[U, G], \varphi) \in sub(\rho, \varphi_0)$. Therefore, we have $(\rho[U, G], \neg\varphi) \in sub^+(\rho, \varphi_0)$.

4. Rule $(\checkmark_1)$: If $(\rho[U, G], \checkmark) \in sub^+(\epsilon, \varphi_0)$ then $(\rho[U, G], \checkmark) \in sub(\epsilon, \varphi_0)$. Let $(\psi, a, \chi) \in U$. We have $(\rho[(\psi, a, \chi), G], \checkmark) \in sub(\epsilon, \varphi_0)$. Since (ρ, ψ), $(\rho, \chi) \in sub(\rho[(\psi, a, \chi), G], \checkmark)$, we have (ρ, ψ), $(\rho, \chi) \in sub(\epsilon, \varphi_0)$. It follows by Proposition 6.4.2 that $(\rho, \checkmark) \in sub(\rho, \psi)$, thus we also have $(\rho, \checkmark) \in sub(\rho, \varphi_0)$.

5. Rule $(\checkmark_2)$: It follows by Proposition 6.4.2 that $(\epsilon, \checkmark), (\epsilon, \mathsf{X}) \in sub(\epsilon, \varphi_0)$.

6. Rule (X_1): If $(\rho[(\psi, a, \chi), G], \mathsf{X}) \in sub^+(\epsilon, \varphi_0)$, then $(\rho[(\psi, a, \chi), G], \mathsf{X}) \in sub(\epsilon, \varphi_0)$. Since $(\rho, \psi), (\rho, \chi) \in sub(\rho[(\psi, a, \chi), G], \mathsf{X})$, we have $(\rho, \psi), (\rho, \chi) \in sub(\epsilon, \varphi_0)$. Therefore, we have $(\rho, \neg\psi)$, $(\rho, \neg\chi) \in sub^+(\epsilon, \varphi_0)$. It follows by Proposition 6.4.2 that $(\rho, \mathsf{X}) \in sub(\rho, \psi)$, thus we also have $(\rho, \mathsf{X}) \in sub(\rho, \varphi_0)$.

$\Box$

Proposition 6.4.4 *Let T be a tableau for φ_0, and let b be a branch of T. If σ is a label present in b, then there are at most k labels present in b with the form of σan for some $n \in \mathbb{N}$, where $k = |sub^+(\epsilon, \varphi_0)|$.*

PROOF It follows by Definition 6.3.5 that each label σan present in b is generated by applying the rule $(\neg\Box_a)$ to a labelled term $(\sigma, \rho, \neg\Box_a\varphi) \in b$. According to Proposition 6.4.3, there are at most k terms labelled with σ in b. Therefore, there are at most k labels present in b with the form of σan for some $n \in \mathbb{N}$. $\Box$

Definition 6.4.5 (Length of label) *The length of a label σ, denoted by $|\sigma|$, is defined by induction on σ: $|n| = 0$; $|\sigma an| = |\sigma| + 1$.*

Proposition 6.4.6 *Let T be a tableau for φ_0 and b be a branch of T. If $(\sigma, \rho, x) \in b$ then $|\sigma| \leq l(\varphi_0) - l(\rho, x)$.*

PROOF Following Definition 6.3.5, the proof is by induction on the process of construction of T. For the initial tableau $\{\{(0, \epsilon, \varphi_0)\}\}$, it is obvious. Next we will show that this property is preserved by all the tableau rules. The cases of the rules $(\neg\neg)$, $(\neg\wedge)$, $(\wedge)$ and $(\checkmark_2)$ are obvious; we will restrict our attention to the other rules.

1. Rule $(\neg\Box_a)$: If $|\sigma| \leq l(\varphi_0) - l(\rho, \neg\Box_a\varphi)$, we have $l(\varphi_0) - l(\rho, \neg\Box_a\varphi) \leq l(\varphi_0) - l(\rho|_a, \neg\varphi) - 1$ because $l(\rho, \neg\Box_a\varphi) \geq l(\rho|_a, \neg\varphi) + 1$. Thus we have $|\sigma| \leq l(\varphi_0) - l(\rho|_a, \neg\varphi) - 1$. It follows that $|\sigma an| \leq l(\varphi_0) - l(\rho|_a, \neg\varphi)$. What is more, since $l(\rho, \neg\Box_a\varphi) \geq l(\rho|_a, \checkmark) + 1$, we have $l(\varphi_0) - l(\rho, \neg\Box_a\varphi) \leq l(\varphi_0) - l(\rho|_a, \checkmark) - 1$. It follows $|\sigma| \leq l(\varphi_0) - l(\rho|_a, \checkmark) - 1$. Thus we have $|\sigma an| \leq l(\varphi_0) - l(\rho|_a, \checkmark)$.

2. Rule $(\Box_a)$: Suppose $|\sigma| \leq l(\varphi_0) - l(\rho, \Box_a\varphi)$, we have $l(\varphi_0) - l(\rho, \Box_a\varphi) \leq l(\varphi_0) - l(\rho|_a, \varphi) - 1$ because $l(\rho, \Box_a\varphi) \geq l(\rho|_a, \varphi) + 1$. Therefore, we have $|\sigma an| \leq l(\varphi_0) - l(\rho|_a, \varphi)$. What is more, since $l(\rho|_a, \varphi) \geq l(\rho|_a, \mathsf{X})$, we have $|\sigma an| \leq l(\varphi_0) - l(\rho|_a, \mathsf{X})$.

3. Rule $(\neg[U, G])$: If $|\sigma| \leq l(\varphi_0) - l(\rho, \neg[U, G]\varphi)$, we have $l(\varphi_0) - l(\rho, \neg[U, G]\varphi) \leq l(\varphi_0) - l(\rho[U, G], \neg\varphi)$ because $l(\rho, \neg[U, G]\varphi) = l(\rho[U, G], \neg\varphi) + 1$. Therefore, we have $|\sigma| \leq l(\varphi_0) - l(\rho[U, G], \neg\varphi)$.

4. Rule $([U, G])$: Since $l(\rho, [U, G]\varphi) = l(\rho[U, G], \varphi) + 1$, if $|\sigma| \leq l(\varphi_0) - l(\rho, [U, G]\varphi)$, we have $|\sigma| \leq l(\varphi_0) - l(\rho[U, G], \varphi)$.

5. Rule $(\checkmark_1)$: Assume $|\sigma an| \leq l(\varphi_0) - l(\rho[U, G], \checkmark)$. Since $l(\rho[U, G], \checkmark) \geq l(\rho, \checkmark)$, it follows that $|\sigma an| \leq l(\varphi_0) - l(\rho, \checkmark)$. Let $(\psi, a, \chi) \in U$. We have $l(\rho[U, G], \checkmark) \geq l(\rho, \psi) - 1$ because $l(\rho[U, G], \checkmark) \geq l(\rho, \psi)$. It follows that $l(\varphi_0) - l(\rho[U, G], \checkmark) \leq l(\varphi_0) - l(\rho, \psi) + 1$. Thus we have $|\sigma an| \leq l(\varphi_0) - l(\rho, \psi) + 1$. It follows that $|\sigma| \leq l(\varphi_0) - l(\rho, \psi)$.

   What is more, since $l(\rho[U, G], \checkmark) \geq l(\rho, \chi)$, it follows that $l(\varphi_0) - l(\rho[U, G], \checkmark) \leq l(\varphi_0) - l(\rho, \chi)$. Thus we have $|\sigma an| \leq l(\varphi_0) - l(\rho, \chi)$.

6. Rule (X_1): Assume that $|\sigma an| \leq l(\varphi_0) - l(\rho[(\psi, a, \chi), G], \mathsf{X})$. Since we have $l(\rho[(\psi, a, \chi), G], \mathsf{X}) \geq l(\rho, \mathsf{X})$, it follows that $|\sigma an| \leq l(\varphi_0) - l(\rho, \mathsf{X})$.

   What is more, since $l(\rho[(\psi, a, \chi), G], \mathsf{X}) \geq l(\rho, \neg\psi)$ and $l(\rho[(\psi, a, \chi), G], \mathsf{X}) \geq l(\rho, \neg\chi)$, it follows that $l(\varphi_0) - l(\rho[(\psi, a, \chi), G], \mathsf{X}) \leq l(\varphi_0) - l(\rho, \neg\psi)$ and $l(\varphi_0) - l(\rho[(\psi, a, \chi), G], \mathsf{X}) \leq l(\varphi_0) - l(\rho, \neg\chi)$. Therefore, we have $|\sigma an| \leq l(\varphi_0) - l(\rho, \neg\psi)$ and $|\sigma an| \leq l(\varphi_0) - l(\rho, \neg\psi)$. Since $|\sigma| \leq |\sigma an|$, it is obvious $|\sigma| \leq l(\varphi_0) - l(\rho, \neg\psi)$.

7. Rule (X_2): Assume $|\sigma an| \leq l(\varphi_0) - l(\rho[U, G], \mathsf{X})$ and $|U| \geq 2$. Suppose that $(\psi, a', \chi) \in U$, we have $l(\rho[U, G], \mathsf{X}) \geq l(\rho[(\psi, a', \chi), G], \mathsf{X})$. It follows that $l(\varphi_0) - l(\rho[U, G], \mathsf{X}) \leq l(\varphi_0) - l(\rho[(\psi, a', \chi), G], \mathsf{X})$. Thus we have $|\sigma an| \leq l(\varphi_0) - l(\rho[(\psi, a', \chi), G], \mathsf{X})$.

$\Box$

Lemma 6.4.7 (Small model property) *If φ_0 is satisfiable then φ_0 is satisfiable in a model which is bounded by $k^{O(m)}$, where $k = |sub^+(\epsilon, \varphi_0)|$ and $m = l(\varphi_0)$.*

PROOF It follows by Theorem 6.3.10 that all tableaux for φ_0 are open. According to the proof of Theorem 6.3.14, we can construct a model $\mathcal{M}^c$ from a saturated branch b such that φ_0 is satisfied in $\mathcal{M}^c$. By the definition of $\mathcal{M}^c$, we know that each state in $\mathcal{M}^c$ is a label present in b. Please note that all labels present in b form a tree. It follows by

Proposition 6.4.4 that each label in the tree has at most k children. It follows by Proposition 6.4.6 that the depth of the tree is bounded by m. Therefore, there are at most $k^{O(m)}$ labels used in b. $\Box$

Theorem 6.4.8 (Decidability) *The problem whether φ_0 is satisfiable is decidable.*

PROOF It follows by Lemma 6.4.7 that we only need to check all the models no bigger than $k^{O(m)}$ where $k = |sub^+(\epsilon, \varphi_0)|$ and $m = l(\varphi_0)$, and this procedure can terminate in finitely many steps. $\Box$

6.5 Conclusion

This chapter presented the theory PAUL of Private Arrow Update Logic, which extends the arrow update of AUL with a relativized subgroup of agents. Public, private and semi-private announcements can be modeled in this framework. PAUL still is a particular case of GAUL, since some information change, like cheating, cannot be modeled in PAUL. This chapter also provided a sound and complete tableau method of PAUL and showed that PAUL is decidable.

In Kooi and Renne (2011a), an axiomatic theory for AUL is provided by the reduction axiom method of DEL (see also Baltag and Moss (2004); Baltag et al. (1998); Gerbrandy (1999); Plaza (2007); van Benthem et al. (2006); van Ditmarsch et al. (2007)). Similarly, we can have the PAUL version reduction axioms. Especially, the reduction axioms of $\Box_a$ formula are as below.

$$[U, G]\Box_a\varphi \leftrightarrow \bigwedge_{(\psi,a,\chi)\in U} (\psi \to \Box_a(\chi \to [U, G]\varphi)) \qquad a \in G$$

$$[U, G]\Box_a\varphi \leftrightarrow \Box_a\varphi \qquad a \notin G$$

Propositions 6.2.5 and 6.2.6 have shown that these two formulas are valid. Therefore, we can have an axiomatic theory for PAUL. The only difference is that there is a reduction axiom for the composition of two updates in Kooi and Renne (2011a), but we cannot generally combine two updates $[U_1, G_1][U_2, G_2]$ here, because it might be the case that $G_1 \neq G_2$. However, this will not be a problem. It is shown in Wang and Cao (2013) that the reduction axiom of composing two update operators is not necessary. Without the composition axiom, we can still complete the reduction by defining the reduction function as $r([U_1, G_1][U_2, G_2]\varphi) = r([U_1, G_1]r([U_2, G_2]\varphi))$. Therefore, the reduction axiom method indeed works for PAUL. Since most of the proofs are similar to those for AUL in Kooi and Renne (2011a), we do not prove it in this chapter.

For future research, we can try to give an optimal algorithm for the satisfiability problem of PAUL by taking a depth-first search strategy on the tableau method. Since each AUL formula can be equivalently translated into a PAUL formula by replacing the update $[U]$ by $[U, \mathbf{Agt}]$, the tableau method presented in this chapter can apply to AUL. Therefore, the optimal algorithm for PAUL (if there is one) will also be an algorithm for AUL and might also be optimal.

Another direction for future research is to apply Arrow Update Logic to knowing how. The main feature of Arrow Update Logic is that it updates information but does not eliminate states. This makes it more suitable for modeling information update in knowing how. For example, a doctor may not know how to treat a patient since the only two available medicines a and b may cause some very bad side-effect. That is, there is an a-arrow and a b-arrow from the current state to the bad side-effect state. If the information is updated, for example, a new scientific discovery shows that a will not cause the bad effect, then the doctor should know how to treat the patient. This kind of information update will eliminate arrows but not states.

Chapter 7

Conclusion

This thesis investigates planning under uncertainty from a logical point of view. In artificial intelligence, conformant planning, which is the simplest version of planning under uncertainty, is to find a linear sequence of actions to achieve goals, where the agent is uncertain about his situation. Conformant probabilistic planning extends conformant planning by extending the uncertainties with probability distributions. Contingent planning generalizes the solution of conformant planning to be a strategy, a partial function from belief spaces to actions. Chapter 2 proposes a logical framework to capture how the probability distribution (which stands for the agent's uncertainty) updates along the execution of a plan in conformant probabilistic planning. Chapter 3 and 4 build knowing-how logics by developing the idea of interpreting knowing how to achieve a goal as having a conformant plan for achieving the goal. Chapter 5 proposes a knowing-how logic by adopting the idea of reducing knowledge-how as having a contingent plan. Chapter 6 extends the theory of arrow update logic (AUL) with the motivation that AUL provides a proper way to model the information change in the knowing-how logic proposed in Chapter 5.

Chapter 2 developed a logical framework for conformant probabilistic planning. This approach differs from existing approaches to conformant probabilistic planning by focussing on a logical language with which to specify plans and goals. The particular logic also allows for reasoning about the change of the belief state of the agent (which is a probability distribution over states) during the execution of actions. In this logic, we can enrich conformant probabilistic planning by thinking of the goal as a formula, which may be more convenient when we formulate goals that are probabilistic in nature. We provided a complete axiomatization of the logic, which shows it is rather well-behaved for a logic that deals with conformant probabilistic planning. We also proved this logic to be decidable.

Chapter 3 and 4 extended Wang's knowing-how logic, in which knowing how to achieve a goal is interpreted as having a conformant plan. Chapter 3 extended the original binary knowing-how modality to be a ternary modality, which can express that the agent knows how to achieve a goal from a specific position while taking a route that satisfies some constraints. We presented a sound and complete axiomatization for this extended knowing-how logic and showed this logic to be decidable by using the filtration method. Chapter 4 weakens the binary knowing-how logic by interpreting knowing how to achieve a goal as having a weak conformant plan. The demands that a conformant plan puts on the plan may be too strong, in the sense that a plan is not supposed to fail during the execution. A weak conformant plan, on the other hand, is a plan that will always result in the goals when the execution of the plan terminates, even if the execution is not completed. We argued that it is sufficient to say that one knows how to achieve a goal if he (she) has a weak conformant plan for achieving the goal. We presented a sound and complete axiomatization for this weaker logic and showed that this logic to be decidable.

Inspired by contingent planning, Chapter 5 extended epistemic logic with a unary knowing-how modality which is interpreted as having a strategy for achieving goals. It is called strategically knowing-how logic. In this logic framework, we investigated the interactions between knowledge-that and knowledge-how. We also presented a sound and complete axiomatization of this logic and showed this logic to be decidable. To reason about information change due to announcements in the strategically knowing-how logic, Chapter 6 investigated arrow update logic. The standard arrow update logic cannot describe information change caused by private events. Chapter 6 extended it to deal with private announcements. We presented a sound and complete tableau system for this extended logic and showed it is decidable.

There is a lot more to explore. Even though in each chapter we have shown the logic to be decidable, we have not touched upon the issue of complexity. How is the model theory of each logic? For instance, what kinds of model classes are definable in the logic, and what kinds of structure are not distinguished by the logic (i.e. what is the "bisimulation" of the logic)? Besides these, there are two that I think are very much worth while.

The first one is to cast all the standard AI planning problems into one unified logical framework to facilitate careful comparisons and classification. As it is mentioned in Chapter 1.2, EPDL can apply for conformant planning. Moreover, the language of EPDL is powerful enough to express contingent plans. The logic proposed in Chapter 2 of this thesis can apply for conformant probabilistic planning. If we merge these two logics and generalize the model with partial observability, the logical framework can deal with all kinds of planning problems under uncertainty: conformant planning, conformant probabilistic planning, contingent planning, and contingent planning extended with probabilities. We will then see clearly how the form of the goal formula, the constructor of the plan, and the observational ability matter in the theoretical and practical complexity of planning, which is in line with the research pioneered in Bäckström and Jonsson (2011).

The other is to consider knowing how in multi-agent settings and to model the group notions of knowing how. One of the main reasons that contribute to epistemic logic's great success is that it is very useful when applied to situations involving more than one agent. It can model the group notions of propositional knowledge, such as common (propositional) knowledge and distributed (propositional) knowledge. Similarly, there are also such group notions of knowledge-how. For example, if you know how to reach s_2 from s_1 and I know how to reach s_3 from s_2, then we two together will know how to reach s_3 from s_1. This is a distributed knowledge-how. Distributed knowledge-how can be viewed as the ability that a "highly skilled person" in a group has, one who can do all the actions that each member of the group can do. Common knowledge-how of a group can be viewed as the ability that every member of the group has.

Appendix A

Logical background

A.1 Epistemic logic

Epistemic logic proposed by Hintikka (1962) is a modal logic concerned with reasoning about knowledge expressed by *knowing that* p, where p is a proposition. This subsection will introduce the language, the semantics and the proof system $\mathbb{S}5$ of epistemic logic. Moreover, we will introduce some basic logical concepts, such as validity, consistency, soundness, completeness.

Let $\mathbf{P}$ be a countable set of proposition letters.

Definition A.1.1 (Language) *The language of epistemic logic is defined as below.*

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

where $p \in \mathbf{P}$. We use the following abbreviations: $\bot := \neg\top$, $\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)$, $\varphi \to \psi := \neg\varphi \vee \psi$.

Intuitively, the formula $\mathcal{K}\varphi$ means that the agent knows that φ holds. The language of epistemic logic is interpreted on models defined below.

Definition A.1.2 (Models) *A model is a triple $\mathcal{M} = \langle S, \sim, V\rangle$ where*

- *$S \neq \emptyset$ is a set of states;*
- *$\sim$ is an equivalence relation on S;*
- *$V : S \to 2^{\mathbf{P}}$ is a valuation function.*

For each $s \in S$, $(\mathcal{M}, s)$ is called a pointed model.

Definition A.1.3 (Semantics) *Let $(\mathcal{M}, s)$ be a pointed model. A formula φ being truth in $\mathcal{M}, s$, denoted as $\mathcal{M}, s \vDash \varphi$, is defined as below.*

$$\begin{array}{lcl}
\mathcal{M}, s \vDash \top & & always \\
\mathcal{M}, s \vDash p & \iff & p \in V(s). \\
\mathcal{M}, s \vDash \neg\varphi & \iff & \mathcal{M}, s \nvDash \varphi. \\
\mathcal{M}, s \vDash (\varphi \wedge \psi) & \iff & \mathcal{M}, s \vDash \varphi \text{ and } \mathcal{M}, s \vDash \psi. \\
\mathcal{M}, s \vDash \mathcal{K}\varphi & \iff & \mathcal{M}, t \vDash \varphi \text{ for each } t \text{ with } s \sim t
\end{array}$$

Next, we introduce the concept of validity, which is a crucial notion of the semantics of a logic framework.

Definition A.1.4 (Semantic consequence) *Let Γ be a set of formulas. We write "$\mathcal{M}, s \vDash \psi$ for all $\psi \in \Gamma$" as $\mathcal{M}, s \vDash \Gamma$. We say φ is a* semantic consequence *of Γ, denoted as $\Gamma \vDash \varphi$, if $\mathcal{M}, s \vDash \Gamma$ implies $\mathcal{M}, s \vDash \varphi$ for each pointed model $(\mathcal{M}, s)$.*

We say a formula φ is valid*, denoted as $\vDash \varphi$, if $\emptyset \vDash \varphi$. In other words, φ is valid if $\mathcal{M}, s \vDash \varphi$ for each pointed model $(\mathcal{M}, s)$. We say φ is* satisfiable *if there exists $(\mathcal{M}, s)$ such that $\mathcal{M}, s \vDash \varphi$.*

The following is the deductive system of epistemic logic, which is known as $\mathbb{S}5$.

Definition A.1.5 (Deductive system $\mathbb{S}5$) *The axioms and rules shown in Table A.1 constitute the deductive system $\mathbb{S}5$.*

	AXIOMS
	All instances of propositional tautologies
`DISTK`	$\mathcal{K}(\varphi \to \psi) \to (\mathcal{K}\varphi \to \mathcal{K}\psi)$
`T`	$\mathcal{K}\varphi \to \varphi$
`4`	$\mathcal{K}\varphi \to \mathcal{K}\mathcal{K}\varphi$
`5`	$\neg\mathcal{K}\varphi \to \mathcal{K}\neg\mathcal{K}\varphi$
	RULES
MP	From $\varphi \to \psi$ and φ, infer ψ
GEN	From φ, infer $\mathcal{K}\varphi$

Table A.1: System $\mathbb{S}5$

Next, we introduce the concept of derivation, which is a core notion of the deductive system of a logic framework.

Definition A.1.6 (Derivation) *A finite sequence of formulas $\varphi_1 \cdots \varphi_n$ is called a* derivation *if each φ_i ($1 \leq i \leq n$) is either an instance of an axiom of $\mathbb{S}5$ or following from the preceding formulas in the sequence by a rule of $\mathbb{S}5$. We say φ is derivable in $\mathbb{S}5$, written as $\mathbb{S}5 \vdash \varphi$ or sometimes just $\vdash \varphi$, if there is a derivation $\varphi_1 \cdots \varphi_n \varphi$. Otherwise, we say φ is not derivable in $\mathbb{S}5$ (written as $\nvdash \varphi$).*

Let Γ be a set of formulas. We say φ is derivable from Γ (written as $\Gamma \vdash \varphi$) if there are finite formulas $\varphi_1, \cdots, \varphi_n \in \Gamma$ such that $\vdash (\varphi_1 \wedge \cdots \wedge \varphi_n) \to \varphi$. Otherwise, we say φ is not derivable from Γ (written as $\Gamma \nvdash \varphi$). We say Γ is $\mathbb{S}5$-consistent *(or just consistent) if $\Gamma \nvdash \bot$. Otherwise, we say Γ is inconsistent.*

The following concepts, soundness and completeness, connect the deductive system and the semantics. If the deductive system is sound and complete, it means the formulas derived in the deductive system coincide with the formulas valid to the semantics.

Definition A.1.7 (Soundness and completeness) *We say $\mathbb{S}5$ is* sound *with respect to the semantics if $\vdash \varphi$ implies $\vDash \varphi$ for each formula φ.*

We say $\mathbb{S}5$ is complete *with respect to the semantics if $\vDash \varphi$ implies $\vdash \varphi$ for each formula φ. We say $\mathbb{S}5$ is* strongly complete *if $\Gamma \vDash \varphi$ implies $\Gamma \vdash \varphi$.*

Theorem A.1.8 $\mathbb{S}5$ *is sound and strongly complete with respect to the semantics of Definition A.1.3.*

A.2 Probabilistic dynamic epistemic logic

Probabilistic dynamic epistemic logic (PDEL) proposed by Kooi (2003) is a combination of the public announcement logic Plaza (2007) (which is a simple and intuitive kind of dynamic epistemic logic) and the probabilistic logic by Fagin et al. (1990), which focuses on reasoning about probability, information, and information change and takes higher order information into account. Just like the probabilistic logic, PDEL is an extension of the linear inequality logic. This subsection will introduce the language, the semantics, and the deductive system of PDEL. Before that, we first introduce the linear inequality logic.

A.2.1 Linear inequality logic

The linear inequality logic proposed in Fagin et al. (1990) is a logic for reasoning about linear inequalities. Fagin et al. introduced the linear inequality logic in their paper about probabilistic logic because the atomic probabilistic formula is a linear inequality. The probabilistic logic proposed in Fagin et al. (1990) is an extension of linear inequality logic. In this part, we first introduce the language, the semantics, and the deductive system of linear inequality logic, and second, we introduce some important properties (on solvability and decidability) of linear inequalities.

Let X be a countable set of variables.

Definition A.2.1 (Language) *The language is defined as below.*

$$\varphi ::= a_0x_0 + \cdots + a_nx_n \geq q \mid \neg\varphi \mid (\varphi \wedge \varphi)$$

where $n \in \mathbb{N}$, $x_i \in X$ *and* $a_i, q \in \mathbb{Q}$ *for all* $0 \leq i \leq n$. *We call formulas of the forms* $a_0x_0 + \cdots + a_nx_n \geq q$ *or* $\neg(a_0x_0 + \cdots + a_nx_n \geq q)$ linear inequalities.

Please note that $\neg(a_0x_0 + \cdots + a_nx_n \geq q)$ can also be written as $a_0x_0 + \cdots + a_nx_n < q$ or $(-1)a_0x_0 + \cdots + (-1)a_nx_n > q$.

Definition A.2.2 (Model) *A model A is an assignment function that assigns a real number to every variable $x \in X$.*

Definition A.2.3 (Semantics) *The satisfaction relation between an assignment and a formula is defined as below.*

$$\begin{array}{lcl} A \vDash a_0x_0 + \cdots + a_nx_n \geq q & \iff & a_0A(x_0) + \cdots + a_nA(x_n) \geq q. \\ A \vDash \neg\varphi & \iff & A \nvDash \varphi. \\ A \vDash (\varphi \wedge \psi) & \iff & A \vDash \varphi \text{ and } A \vDash \psi. \end{array}$$

We say φ is satisfiable/solvable if there exists an assignment A such that $A \vDash \varphi$.

Definition A.2.4 (Deductive system) *The axioms and rules shown in Table A.2 constitute the deductive system $\mathbb{SLIL}$.*

	AXIOMS
	All instances of propositional tautologies
Linear inequality axioms	
Identity	$x \geq x$
0 terms	$\sum_{i=1}^{n} q_i x_i \geq q \leftrightarrow \sum_{i=1}^{n} q_i x_i + 0x' \geq q$
Permutation	$\sum_{i=1}^{n} q_i x_i \geq q \rightarrow \sum_{i=1}^{n} q_{k_i} x_{k_i} \geq q$ where $k_1, \cdots k_n$ is a permutation of $1, \cdots n$.
Addition	$(\sum_{i=1}^{n} q_i x_i \geq q) \wedge (\sum_{i=1}^{n} q'_i x_i \geq q') \rightarrow \sum_{i=1}^{n} (q_i + q'_i) x_i \geq q + q'$
Multiplication	$\sum_{i=1}^{n} q_i x_i \geq q \leftrightarrow \sum_{i=1}^{n} d q_i x_i \geq dq$ where d is positive rational.
Dichotomy	$(x \geq q) \vee (x \leq q)$
Monotonicity	$(x \geq q) \rightarrow (x > q')$ where $q > q'$
	RULE
MP	From $\varphi \rightarrow \psi$ and φ, infer ψ

Table A.2: System $\mathbb{SLIL}$

The following theorem is proved in Fagin et al. (1990).

Theorem A.2.5 *$\mathbb{SLIL}$ is sound and complete with respect to the semantics of Definition A.2.3.*

Next, we introduce two properties of linear inequalities on the solvability and the decidability of linear inequalities. Before that, we first introduce the following auxiliary notion.

Definition A.2.6 (Legal linear combination Kuhn (1956)) *Let S be a set of linear inequalities, which can be written as*

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &> a_i && (i = 1, \cdots, p) \\ b_{j1}x_1 + \cdots + b_{jn}x_n &\geq b_j && (j = 1, \cdots, q) \end{aligned}$$

where a_{ik}, a_i, b_{jk} and b_j $(i = 1, \cdots, p; j = 1, \cdots, q; k = 1, \cdots, n)$ are given rational numbers. A multiplier scheme of S is formed with non-negative multipliers at the left and the sum below:

$$\begin{array}{ll} u_0 \geq 0: & 0x_1 + \cdots + 0x_n > -1 \\ u_1 \geq 0: & a_{11}x_1 + \cdots + a_{1n}x_n > a_1 \\ \multicolumn{2}{l}{\dotfill} \\ u_p \geq 0: & a_{p1}x_1 + \cdots + a_{pn}x_n > a_p \\ v_1 \geq 0: & b_{11}x_1 + \cdots + b_{1n}x_n \geq b_1 \\ \multicolumn{2}{l}{\dotfill} \\ v_q \geq 0: & b_{q1}x_1 + \cdots + b_{qn}x_n \geq b_q \\ \hline & d_1x_1 + \cdots + d_nx_n > d. \end{array}$$

The coefficients of the sum are calculated to be:

$$\begin{aligned} d_1 &= u_1 a_{11} + \cdots + u_p a_{p1} + v_1 b_{11} + \cdots v_q b_{q1} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ d_n &= u_1 a_{1n} + \cdots + u_p a_{pn} + v_1 b_{1n} + \cdots v_q b_{qn} \\ d &= -u_0 + u_1 a_1 + \cdots + u_p a_p + v_1 b_1 + \cdots v_q b_q. \end{aligned}$$

An inequality,

$$d_1 x_1 + \cdots + d_n x_n > d,$$

that is formed in this manner from S is called a legal linear combination *of the inequalities of S provided that some u_i is positive ($i = 0, 1, \cdots, p$).*

The following theorem is proved in Kuhn (1956).

Theorem A.2.7 *If a set of linear inequalities S is not solvable, then the inequality $0x_1 + \cdots + 0x_n > 0$ is a legal linear combination of the inequalities of S.*

The size of a rational number a/b, where a and b are integers and relatively prime, is defined to be the sum of the lengths of a and b, when written in binary. The following theorem is proved in Fagin et al. (1990).

Theorem A.2.8 *If a system of r linear inequalities with integer coefficients each of length at most l has a nonnegative solution, then it has a nonnegative solution with at most r entries positive, and where the size of each member of the solution is $O(rl + r\log(r))$.*

A.2.2 Probabilistic dynamic epistemic logic

In the following, we will introduce the language, the semantics, and the deductive system of PDEL.

Let $\mathbf{P}$ be a countable set of proposition letters and $\mathbf{Agt}$ be a nonempty finite set of agents.

Definition A.2.9 (Language) *The language of PDEL is defined as below.*

$$\varphi ::= \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box_a \varphi \mid [\varphi]\varphi \mid q_1 \mathbf{Pr}_a \varphi_1 + \cdots + q_n \mathbf{Pr}_a \varphi_n \geq q$$

where $p \in \mathbf{P}$ and $a \in \mathbf{Agt}$.

Definition A.2.10 (Model) *A model is a tuple $\langle S, R, P, V\rangle$ where*

- *$S \neq \emptyset$ is a set of states;*
- *$R : \mathbf{Agt} \to 2^{S\times S}$ is a collection of transitions labelled by actions in $\mathbf{Agt}$;*
- *$V : S \to 2^{\mathbf{P}}$ is a valuation function;*
- *$P : (S \times \mathbf{Agt}) \to (S \to [0, 1])$ assigns a probability function to each agent on each state such that for each $a \in \mathbf{Agt}$ and each $s \in S$:*

$$\sum_{t \in S} P(s, a)(t) = 1.$$

.

For each $s \in S$, $(\mathcal{M}, s)$ is a pointed model.

Definition A.2.11 (Semantics) *Given a pointed model $(\mathcal{M}, s)$ and a formula φ, φ being true in $(\mathcal{M}, s)$ (written as $\mathcal{M}, s \vDash \varphi$) is defined by induction on φ:*

$$
\begin{array}{lcl}
\mathcal{M}, s \vDash \top & & \textit{always} \\
\mathcal{M}, s \vDash p & \iff & p \in V(s). \\
\mathcal{M}, s \vDash \neg\varphi & \iff & \mathcal{M}, s \nvDash \varphi. \\
\mathcal{M}, s \vDash (\varphi \wedge \psi) & \iff & \mathcal{M}, s \vDash \varphi \textit{ and } \mathcal{M}, s \vDash \psi. \\
\mathcal{M}, s \vDash \Box_a \varphi & \iff & \mathcal{M}, t \vDash \varphi \textit{ for each } t \textit{ with } (s,t) \in R(a) \\
\mathcal{M}, s \vDash \sum_{i=1}^{n} q_i \mathbf{Pr}_a \varphi_i \geq q & \iff & \sum_{i=1}^{n} q_i P(a,s)[\![\varphi_i]\!] \geq q
\end{array}
$$

where $[\![\varphi]\!] = \{s \in S \mid \mathcal{M}, s \vDash \varphi\}$ and $P(a,s)[\![\varphi_i]\!] = \sum_{t \in [\![\varphi]\!]} P(a,s)(t)$.

Definition A.2.12 (Deductive system $\mathbb{SPDEL}$) *The axioms and rules shown in Table A.3 constitute the deductive system $\mathbb{SPDEL}$.*

	AXIOMS
	All instances of propositional tautologies
	All instances of linear inequality axioms
$\Box_a$-distribution	$\Box_a(\varphi \to \psi) \to (\Box_a\varphi \to \Box_a\psi)$
Update axioms	
$[\varphi]$-distribution	$[\varphi](\psi \to \chi) \to ([\varphi]\psi \to [\varphi]\chi)$
Atomic Permanence	$[\varphi]p \leftrightarrow (\varphi \to p)$
Functionality	$\neg[\varphi]\psi \leftrightarrow [\varphi]\neg\psi$
$\Box_a$-update	$[\varphi]\Box_a\psi \leftrightarrow \Box_a(\varphi \to [\varphi]\psi)$
Probability update1	$\mathbf{Pr}_a\varphi > 0 \to ([\varphi] \sum_{i=1}^{n} q_i \mathbf{Pr}_a\varphi_i \geq q \leftrightarrow \sum_{i=1}^{n} q_i \mathbf{Pr}_a(\varphi \wedge [\varphi]\varphi_i) \geq q\mathbf{Pr}_a\varphi)$
Probability update2	$\mathbf{Pr}_a\varphi = 0 \to ([\varphi] \sum_{i=1}^{n} q_i \mathbf{Pr}_a\varphi_i \geq q \leftrightarrow \sum_{i=1}^{n} q_i \mathbf{Pr}_a([\varphi]\varphi_i) \geq q)$
Probability axioms	
Nonnegativity	$\mathbf{Pr}_a\varphi \geq 0$
Probability of truth	$\mathbf{Pr}_a\top = 1$
Additivity	$\mathbf{Pr}_a(\varphi \wedge \psi) + \mathbf{Pr}_a(\varphi \wedge \neg\psi) = \mathbf{Pr}_a\varphi$
	RULES
MP	From $\varphi \to \psi$ and φ, infer ψ
$\Box_a$-GEN	From φ, infer $\Box_a\varphi$
$[\varphi]$-GEN	From ψ, infer $[\varphi]\psi$
Equivalence	From $\varphi \leftrightarrow \psi$, infer $\mathbf{Pr}_a\varphi = \mathbf{Pr}_a\psi$

Table A.3: System $\mathbb{SPDEL}$

The following theorem is proved in Kooi (2003).

Theorem A.2.13 $\mathbb{SPDEL}$ *is sound and complete with respect to the semantics of Definition A.2.11.*

A.3 Knowing-how Logic

Wang (2015a) proposed a modal logic for reasoning about knowledge expressed by *knowing how to achieve that* p. In this thesis, we will call it *knowing-how logic* (KHL). This subsection will introduce the language, the semantics, and the deductive system of KHL.

Let $\mathbf{P}$ be a countable set of proposition letters.

Definition A.3.1 (Language) *The language of KHL is defined as below.*

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

where $p \in \mathbf{P}$.

The language of KHL is interpreted on models defined below.

Definition A.3.2 (Model) *A model is a quadruple* $\mathcal{M} = \langle S, Act, R, V\rangle$ *where*

- $S \neq \emptyset$ *is a set of states;*
- Act *is a set of actions;*
- $R : Act \to 2^{S\times S}$ *is a collection of transitions labelled by actions in* Act*;*
- $V : S \to 2^{\mathbf{P}}$ *is a valuation function.*

For each $s \in S$, $(\mathcal{M}, s)$ *is called a pointed model.*

Please note that the action set is part of the model. We use Act^* to denote all the finite sequences of members of Act. We write $s \xrightarrow{a} t$ if $(s,t) \in R(a)$. For a sequence $\sigma = a_1 \dots a_n \in Act^*$, we write $s \xrightarrow{\sigma} t$ if there exist $s_2 \dots s_n$ such that $s \xrightarrow{a_1} s_2 \xrightarrow{a_2} \cdots \xrightarrow{a_{n-1}} s_n \xrightarrow{a_n} t$. Note that σ can be the empty sequence ϵ (when $n = 0$), and we set $s \xrightarrow{\epsilon} s$ for any s. Let σ_k be the initial segment of σ up to a_k for $k \leq |\sigma|$. In particular let $\sigma_0 = \epsilon$.

Before defining the semantics of KHL, we first introduce the following auxiliary notion.

Definition A.3.3 (Strongly executable) *We say* $\sigma = a_1 \cdots a_n$ *is strongly executable at* s' *if for each* $0 \leq k < n$*:* $s' \xrightarrow{\sigma_k} t$ *implies that* t *has at least one* a_{k+1}*-successor.*

Definition A.3.4 (Semantics) *Let* $(\mathcal{M}, s)$ *be a pointed model. A formula* φ *being true in* $\mathcal{M}, s$*, denoted as* $\mathcal{M}, s \vDash \varphi$*, is defined as below.*

$$\begin{array}{lll}
\mathcal{M}, s \vDash \top & & \textit{always} \\
\mathcal{M}, s \vDash p & \iff & p \in V(s). \\
\mathcal{M}, s \vDash \neg\varphi & \iff & \mathcal{M}, s \nvDash \varphi. \\
\mathcal{M}, s \vDash (\varphi \wedge \psi) & \iff & \mathcal{M}, s \vDash \varphi \textit{ and } \mathcal{M}, s \vDash \psi. \\
\mathcal{M}, s \vDash \mathcal{K}h(\psi, \varphi) & \iff & \textit{there exists a } \sigma \in Act^* \textit{ such that for all } s' \in [\![\psi]\!] : \\
 & & \sigma \textit{ is strongly executable at } s' \textit{ and} \\
 & & \mathcal{M}, t \vDash \varphi \textit{ for all } t \textit{ with } s' \xrightarrow{\sigma} t
\end{array}$$

where $[\![\psi]\!] = \{s' \in S \mid \mathcal{M}, s' \vDash \psi\}$.

By the semantics above, we can see that the modality $\mathcal{K}h$ is a universal operator. Furthermore, it is easy to check that

$$\mathcal{M}, s \vDash \mathcal{K}h(\neg\varphi, \bot) \iff \mathcal{M}, s' \vDash \varphi \text{ for each } s' \in S.$$

Therefore, the formula $\mathcal{K}h(\neg\varphi, \bot)$ is also written as $\mathcal{U}\varphi$, which means that φ is true in each state of the model $\mathcal{M}$.

Definition A.3.5 (Deductive system) *The axioms and rules shown in Table A.4 constitute the deductive system* $\mathbb{SKH}$.

	AXIOMS
`TAUT`	all tautologies of propositional logic
`DISTU`	$\mathcal{U}p \wedge \mathcal{U}(p \to q) \to \mathcal{U}q$
`TU`	$\mathcal{U}p \to p$
`4KU`	$\mathcal{K}h(p,q) \to \mathcal{U}\mathcal{K}h(p,q)$
`5KU`	$\neg\mathcal{K}h(p,q) \to \mathcal{U}\neg\mathcal{K}h(p,q)$
`EMPKh`	$\mathcal{U}(p \to q) \to \mathcal{K}h(p,q)$
`COMPKh`	$(\mathcal{K}h(p,r) \wedge \mathcal{K}h(r,q)) \to \mathcal{K}h(p,q)$
	RULES
MP	From $\varphi \to \psi$ and φ, infer ψ
GEN	From φ, infer $\mathcal{U}\varphi$
SUB	From φ, infer $\varphi[\psi/p]$

Table A.4: System $\mathbb{SKH}$

The following theorem is proved in Wang (2015a).

Theorem A.3.6 $\mathbb{SKH}$ *is sound and strongly complete with respect to the semantics of Definition A.3.4.*

A.4 Arrow update logic

Arrow update logic proposed in Kooi and Renne (2011a) is a dynamic epistemic logic concerned with reasoning about information change by eliminating epistemic accesses. This subsection will introduce the language, the semantics, and the deductive system of arrow update logic.

Let $\mathbf{P}$ be a countable set of proposition letters and $\mathbf{Agt}$ be a nonempty finite set of agents.

Definition A.4.1 (Language) *The language of arrow update logic is defined as below.*

$$\begin{aligned}\varphi &::= \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box_a\varphi \mid [U]\varphi \\ U &::= (\varphi, a, \varphi) \mid (\varphi, a, \varphi), U\end{aligned}$$

where $p \in \mathbf{P}$ *and* $a \in \mathbf{Agt}$.

Intuitively, the formula $\Box_a\varphi$ means that the agent a believes that φ holds, and the formula $[U]\varphi$ means that φ holds after the model is updated with U.

Definition A.4.2 (Model) *A model is a triple $\mathcal{M} = \langle S, R, V\rangle$ where*

- $S \neq \emptyset$ *is a set of states;*
- $R : \mathbf{Agt} \to 2^{S\times S}$ *is a collection of transitions labelled by actions in* $\mathbf{Agt}$*;*
- $V : S \to 2^{\mathbf{P}}$ *is a valuation function.*

For each $s \in S$, $(\mathcal{M}, s)$ is called a pointed model.

Definition A.4.3 (Semantics) *Given a model $\mathcal{M} = \langle S, R, V\rangle$, let s be a state in S. A formula φ being truth in $\mathcal{M}, s$, denoted as $\mathcal{M}, s \vDash \varphi$, is defined as below.*

$$
\begin{array}{lcl}
\mathcal{M}, s \vDash \top & & always \\
\mathcal{M}, s \vDash p & \iff & p \in V(s). \\
\mathcal{M}, s \vDash \neg\varphi & \iff & \mathcal{M}, s \nvDash \varphi. \\
\mathcal{M}, s \vDash (\varphi \wedge \psi) & \iff & \mathcal{M}, s \vDash \varphi \textit{ and } \mathcal{M}, s \vDash \psi. \\
\mathcal{M}, s \vDash \Box_a\varphi & \iff & \mathcal{M}, t \vDash \varphi \textit{ for each } t \textit{ with } (s,t) \in R(a) \\
\mathcal{M}, s \vDash [U]\varphi & \iff & \mathcal{M} * U, s \vDash \varphi \\
R^{\mathcal{M}*U} & = & \{(w,v) \mid \exists(\psi, a, \chi) \in U : \mathcal{M}, w \vDash \psi \textit{ and } \mathcal{M}, v \vDash \chi\}
\end{array}
$$

where $\mathcal{M} * U = \langle S, R^{\mathcal{M}*U}, V\rangle$.

$\mathcal{M} * U$ is the resulting model after updating $\mathcal{M}$ with U. Please note that the updated model $\mathcal{M} * U$ shares the same domain and the same valuation function with the original model $\mathcal{M}$. The transitions of $\mathcal{M} * U$ is a subset of the transitions of $\mathcal{M}$ because the update U deletes some transitions of $\mathcal{M}$.

Definition A.4.4 (Deductive system $\mathbb{AUL}$) *The axioms and rules shown in Table A.5 constitute the deductive system $\mathbb{AUL}$.*

The following theorem is proved in Kooi and Renne (2011a).

Theorem A.4.5 *$\mathbb{AUL}$ is sound and complete with respect to the semantics of Definition A.4.3.*

	AXIOMS
TAUT	all tautologies of propositional logic
DISTU	$\Box_a p \wedge \Box_a(p \to q) \to \Box_a q$
U1	$[U]p \leftrightarrow p$ for $p \in \mathbf{P} \cup \{\bot, \top\}$
U2	$[U]\neg\varphi \leftrightarrow \neg[U]\varphi$
U3	$[U](\varphi \wedge \psi) \leftrightarrow [U]\varphi \wedge [U]\psi$
U4	$[U]\Box_a\varphi \leftrightarrow \bigwedge_{(\psi,a,\chi)\in U}(\psi \to \Box_a(\chi \to [U]\varphi))$
U5	$[U_1][U_2]\varphi \leftrightarrow [U_1 \circ U_2]\varphi$ where $[U_1 \circ U_2] = \{(\psi \wedge [U_1]\psi', a, \chi \wedge [U_1]\chi') \mid \exists(\psi, a, \chi) \in U_1$ and $\exists(\psi', a, \chi') \in U_2\}$
	RULES
MP	From $\varphi \to \psi$ and φ, infer ψ
GENK	From φ, infer $\Box_a\varphi$
GENU	From φ, infer $[U]\varphi$

Table A.5: System $\mathbb{AUL}$

Bibliography

Ågotnes, T. (2006). Action and knowledge in alternating-time temporal logic. *Synthese*, 149(2):375–407.

Ågotnes, T., Goranko, V., Jamroga, W., and Wooldridge, M. (2015). Knowledge and ability. In van Ditmarsch, H., Halpern, J., van der Hoek, W., and Kooi, B., editors, *Handbook of Epistemic Logic*, chapter 11, pages 543–589. College Publications.

Alur, R., Henzinger, T., and Kupferman, O. (2002). Alternating-time temporal logic. *Journal of the ACM*, 49:672–713.

Andersen, M. B., Bolander, T., and Jensen, M. H. (2012). Conditional epistemic planning. In *Logics in Artificial Intelligence*, pages 94–106. Springer.

Andersen, M. B., Bolander, T., and Jensen, M. H. (2015). Don't plan for the unexpected: Planning based on plausibility models. *Logique et Analyse*, 58.

Aucher, G. (2012). DEL-sequents for regression and epistemic planning. *Journal of Applied Non-Classical Logics*, 22(4):337–367.

Aucher, G. and Bolander, T. (2013). Undecidability in epistemic planning. In *Proceedings of IJCAI '13*, pages 27–33.

Aucher, G. and Schwarzentruber, F. (2013). On the complexity of dynamic epistemic logic. In Schipper, B. C., editor, *Proceedings of the 14th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2013)*, pages 19–28.

Bäckström, C. and Jonsson, P. (2011). All PSPACE-complete planning problems are equal but some are more equal than others. In *SOCS 2011*.

Balbiani, P., van Ditmarsch, H. P., Herzig, A., and Lima, T. D. (2010). Tableaux for public announcement logic. *J. Log. Comput.*, 20(1):55–76.

Baltag, A. and Moss, L. S. (2004). Logics for epistemic programs. *Synthese*, 139:165–224.

Baltag, A., Moss, L. S., and Solecki, S. (1998). The logic of public announcements, common knowledge and private suspicions. In Gilboa, I., editor, *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK VII)*, pages 43–56.

Belardinelli, F. (2014). Reasoning about knowledge and strategies: Epistemic strategy logic. In *Proceedings of SR 2014*, pages 27–33.

Blackburn, P., de Rijke, M., and Venema, Y. (2001). *Modal Logic*. Cambridge University Press.

Bolander, T. and Andersen, M. B. (2011). Epistemic planning for single and multi-agent systems. *Journal of Applied Non-Classical Logics*, 21(1):9–34.

Bolander, T., Jensen, M. H., and Schwarzentruber, F. (2015). Complexity results in epistemic planning. In *IJCAI*, pages 2791–2797.

Broersen, J. and Herzig, A. (2015). Using STIT theory to talk about strategies. In van Benthem, J., Ghosh, S., and Verbrugge, R., editors, *Models of Strategic Reasoning: Logics, Games, and Communities*, pages 137–173. Springer.

Bryce, D., Kambhampati, S., and Smith, D. E. (2008). Sequential monte carlo in reachability heuristics for probabilistic planning. *Artificial Intelligence*, 172(6-7):685–715.

Fagin, R., Halpern, J., Moses, Y., and Vardi, M. (1995). *Reasoning about Knowledge*. MIT Press.

Fagin, R., Halpern, J. Y., and Megiddo, N. (1990). A logic for reasoning about probabilities. *Inf. Comput.*, 87(1-2):78–128.

Fan, J., Wang, Y., and van Ditmarsch, H. (2014). Almost necessary. In *Advances in Modal Logic Vol.10*, pages 178–196.

Fan, J., Wang, Y., and van Ditmarsch, H. (2015). Contingency and knowing whether. *The Review of Symbolic Logic*, 8:75–107.

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

Fervari, R., Herzig, A., Li, Y., and Wang, Y. (2017). Strategically knowing how. In *Proceedings of IJCAI 2017*, pages 1031–1038.

Fitting, M. (1983). *Proof Methods for Modal and Intuitionistic Logics*. Springer.

Geffner, H. and Bonet, B. (2013). *A Concise Introduction to Models and Methods for Automated Planning*. Morgan & Claypool Publishers.

Gerbrandy, J. (1999). *Bisimulations on Planet Kripke*. PhD thesis, University of Amsterdam.

Gerbrandy, J. and Groeneveld, W. (1997). Reasoning about information change. *Journal of Logic, Language and Information*, 6:147–169.

Ghallab, M., Nau, D., and Traverso, P. (2004). *Automated Planning: theory and practice*. Morgan Kaufmann.

Giunchiglia, F. and Traverso, P. (2000). Planning as model checking. In Biundo, S. and Fox, M., editors, *Recent Advances in AI Planning*, pages 1–20. Springer Berlin Heidelberg.

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

Gu, T. and Wang, Y. (2016). "Knowing value" logic as a normal modal logic. In *Advances in Modal Logic Vol.11*, pages 362–381.

Hart, S., Heifetz, A., and Samet, D. (1996). Knowing whether, knowing that, and the cardinality of state spaces. *Journal of Economic Theory*, 70(1):249–256.

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

Herzig, A., Lorini, E., and Walther, D. (2013). Reasoning about actions meets strategic logics. In *Proceedings of LORI 2013*, pages 162–175.

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

Hyafil, N. and Bacchus, F. (2003). Conformant probabilistic planning via CSPs. In *ICAPS*, pages 205–214.

Hyafil, N. and Bacchus, F. (2004). Utilizing structured representations and CSPs in conformant probabilistic planning. In *Proceedings of the 16th European Conference on Artificial Intelligence*, ECAI'04, pages 1033–1034.

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.

Jamroga, W. and van der Hoek, W. (2004). Agents that know how to play. *Fundamenta informaticae*, 63(2-3):185–219.

Jensen, M. H. (2014). *Epistemic and Doxastic Planning*. PhD thesis, Technical University of Denmark.

Kooi, B. (2003). Probabilistic dynamic epistemic logic. *Journal of Logic, Language, and Information*, 12(4):381–408.

Kooi, B. and Renne, B. (2011a). Arrow update logic. *The Review of Symbolic Logic*, 4:536–559.

Kooi, B. and Renne, B. (2011b). Generalized arrow update logic. In Apt, K. R., editor, *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Thirteenth Conference (TARK 2011)*, pages 205–211.

Kuhn, H. W. (1956). Solvability and consistency for linear equations and inequalities. *The American Mathematical Monthly*, 63(4):217–232.

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

Lau, T. and Wang, Y. (2016). Knowing your ability. *The Philosophical Forum*, 47(3-4):415–423.

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. (2015). Tableaux for single-agent epistemic PDL with perfect recall and no miracles. In *Proceedings of LORI '15*, pages 230–242.

Li, Y. (2016). *A Dynamic Epistemic Framework for Conformant Planning (in Chinese)*. PhD thesis, Peking University.

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

Li, Y. and Wang, Y. (2017). Achieving while maintaining: - A logic of knowing how with intermediate constraints. In *Logic and Its Applications - 7th Indian Conference, ICLA 2017, Kanpur, India, January 5-7, 2017, Proceedings*, pages 154–167.

Li, Y., Yu, Q., and Wang, Y. (2017). More for free: A dynamic epistemic framework for conformant planning over transition systems. *Journal of Logic and Computation*. to appear.

Löwe, B., Pacuit, E., and Witzel, A. (2011). DEL planning and some tractable cases. In *LORI 2011*, pages 179–192. Springer.

Massacci, F. (2000). Single step tableaux for modal logics. *Journal of Automated Reasoning*, 24:319–364.

McCarthy, J. (1979). First order theories of individual concepts and propositions. *Machine Intelligence*, 9.:129–147.

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

Moore, R. C. (1985). A formal theory of knowledge and action. In Hobbs, J. R. and Moore, R. C., editors, *Formal Theories of the Commonsense World*. Ablex Publishing Corporation.

Muise, C., Belle, V., Felli, P., McIlraith, S., Miller, T., Pearce, A. R., and Sonenberg, L. (2015). Planning over multi-agent epistemic states: A classical planning approach. In *The 29th AAAI Conference on Artificial Intelligence*, pages 3327–3334.

Pardo, P. and Sadrzadeh, M. (2013). Strong planning in the logics of communication and change. In *Declarative Agent Languages and Technologies X*, pages 37–56. Springer.

Plaza, J. (1989). Logics of public communications. In Emrich, M. L., Pfeifer, M. S., Hadzikadic, M., and Ras, Z. W., editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216.

Plaza, J. (2007). Logics of public communications. *Synthese*, 158:165–179. Reprint of Plaza (1989).

Quine, W. V. (1953). Reference and modality. In *From a Logical Point of View*. Harvard University Press.

Rabiner, L. R. (1990). A tutorial on hidden markov models and selected applications in speech recognition. In Waibel, A. and Lee, K.-F., editors, *Readings in Speech Recognition*, pages 267–296. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Singh, M. P. (1994). *Multiagent Systems: A Theoretical Framework for Intentions, Know-how, and Communications*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Smith, D. E. and Weld, D. S. (1998). Conformant graphplan. In *AAAI 98*, pages 889–896.

Stanley, J. and Williamson, T. (2001). Knowing how. *The Journal of Philosophy*, pages 411–444.

Taig, R. and Brafman, R. I. (2013). Compiling conformant probabilistic planning problems into classical planning. In *ICAPS*, pages 197–205.

van Benthem, J. (2003). Conditional probability meets update logic. *Journal of Logic, Language and Information*, 12(4):409–421.

van Benthem, J., Gerbrandy, J., and Kooi, B. (2009). Dynamic update with probabilities. *Studia Logica*, 93(1):67–96.

van Benthem, J., van Eijck, J., and Kooi, B. (2006). Logics of communication and change. *Information and Computation*, 204:1620–1662.

van der Hoek, W. and Lomuscio, A. (2003). Ignore at your peril - towards a logic for ignorance. In *Proceedings of AAMAS 03*, pages 1148–1149.

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

van der Hoek, W. and Wooldridge, M. (2002). Tractable multiagent planning for epistemic goals. In *Proceedings of AAMAS'02*, pages 1167–1174.

van der Hoek, W. and Wooldridge, M. (2003). Cooperation, knowledge, and time: Alternating-time temporal epistemic logic and its applications. *Studia Logica*, pages 125–157.

van Ditmarsch, H., Halpern, J. Y., van der Hoek, W., and Kooi, B., editors (2015). *Handbook of Epistemic Logic*. College Publications.

van Ditmarsch, H., van der Hoek, W., and Kooi, B. (2007). *Dynamic Epistemic Logic*. Springer.

Wang, Y. (2011). Reasoning about protocol change and knowledge. In Banerjee, M. and Seth, A., editors, *Logic and Its Applications - 4th Indian Conference (ICLA 2011)*, pages 189–203.

Wang, Y. (2015a). A logic of knowing how. In *Proceedings of LORI 2015*, pages 392–405.

Wang, Y. (2015b). Representing imperfect information of procedures with hyper models. In *Proceedings of ICLA 2015*, pages 218–231.

Wang, Y. (2016). A logic of goal-directed knowing how. *Synthese*, pages 1–21.

Wang, Y. (2017). Beyond knowing that: a new generation of epistemic logics. In van Ditmarsch, H. and Sandu, G., editors, *Jaakko Hintikka on Knowledge and Game Theoretical Semantics*. Springer.

Wang, Y. and Cao, Q. (2013). On axiomatizations of public announcement logic. *Synthese*, 190:103–134.

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

Yu, Q., Li, Y., and Wang, Y. (2016). A dynamic epistemic framework for conformant planning. In *Proceedings of TARK'15*, pages 298–318. EPTCS.

Yu, Q., Wen, X., and Liu, Y. (2013). Multi-agent epistemic explanatory diagnosis via reasoning about actions. In *IJCAI 2013*, pages 1183–1190.

Samenvatting[1]

In dit proefschrift wordt vanuit logisch perspectief het maken van plannen en procedurele kennis (weten hoe) onderzocht. Binnen de kunstmatige intelligentie verstaat men onder de term conformant planning het proces van het maken van een plan om een bepaald doel te bereiken. Zo'n plan wordt door een persoon of andere handelende entiteit, een zogeheten agent, uitgevoerd. Een agent die doelgerichte procedurele kennis heeft weet hoe zij haar doel kan bereiken. In dit proefschrift wordt doelgerichte procedurele kennis opgevat als (i) het hebben van een plan en (ii) weten dat het uitvoeren van dat plan het doel zal bereiken.

In Hoofdstuk 2 introduceer ik een logisch raamwerk waarin precies kan worden bijgehouden hoe de overtuigingstoestand van een agent (inclusief probabilistische informatie) verandert tijdens het uitvoeren van een plan. Deze benadering verschilt van bestaande benaderingen voor conformant probabilistic planning doordat het zich richt op een logische taal waarin plannen en doelen gespecificeerd worden.

In hoofdstuk 3 and 4 worden logica's gepresenteerd voor procedurele kennis gebaseerd op het idee dat weten hoe je een doel bereikt neerkomt op het hebben van een conformant plan om dat doel te bereiken. In hoofdstuk 3 wordt een logisch systeem gepresenteerd waarbij niet alleen het doel van het plan centraal staat maar ook de wijze waarop dat doel bereikt wordt, d.w.z. het plan zelf moet aan bepaalde voorwaarden voldoen behalve dat het zijn doel bereikt. In hoofdstuk 4 wordt een logica voor procedurele kennis gepresenteerd waarbij de kennis om een doel te bereiken wordt opgevat als het hebben van een zwak conformant plan. De vereisten die een conformant plan stelt aan het plan kunnen namelijk te sterk zijn, in die zin dat er bij het uitvoeren van een plan niets hoort mis te gaan. Een zwak conformant plan is een plan dat altijd tot een bepaald doel leid als het uitvoeren van het plan ophoudt, zelfs als de uitvoering van het plan nog niet voltooid is.

In hoofdstuk 5 wordt een logisch raamwerk voor procedurele kennis gepresenteerd waarbij weten hoe je een doel bereikt opgevat wordt als het hebben van een contingent plan om een doel te bereiken. Anders dan een conformant plan, wat een lineaire reeks handelingen is, is een contingent plan een partiële functie van overtuigingstoestanden naar handelingen. Binnen dit logische raamwerk onderzoeken we ook de interactie tussen propositionele kennis (weten dat) en procedurele kennis (weten hoe).

Arrow update logic (AUL) is een logisch raamwerk dat geschikt is om redeneringen over informatieverandering te formaliseren. In AUL wordt informatieverandering gemodelleert door de epistemische toegankelijkheidsrelatie over mogelijke werelden te

[1] This is a summary of this thesis in Dutch, and everything here is also explained in the chapter of introduction.

wijzigen. Deze manier van modeleren maakt het een geschikte manier om informatieverandering te modeleren in logica's voor procedurele kennis omdat het domein van een model (de verzameling mogelijke werelden) niet wijzigt in tegenstelling tot andere benaderingen. In AUL is het echter gemeenschappelijke kennis hoe elke agent nieuwe informatie zal verwerken. In hoofdstuk 6 wordt de theorie van AUL uitgebreid in die zin dat de nieuwe informatie beperkt blijft tot een bepaalde groep agents.

About the author

Yanjun Li was born in Hebei, China, on the 18th of February in 1985. He graduated from Hebei University in 2010 and obtained a Bachelor's degree in philosophy with a thesis on Frege's theory of meaning. After that, He joined the Center for Logic, Language and Cognition at the Department of Philosophy, Peking University, first as a Master student later as a PhD student, which concluded with a doctorate in 2016 with the thesis "A Dynamic Epistemic Logic for Conformant Planning". In 2013, he moved to The Netherlands and started working on this thesis to obtain a doctorate at the University of Groningen.