

# A New Modal Framework for Epistemic Logic\*

Yanjing Wang

Department of Philosophy, Peking University, China

y.wang@pku.edu.cn

Recent years witnessed a growing interest in non-standard epistemic logics of knowing whether, knowing how, knowing what, knowing why and so on. The new epistemic modalities introduced in those logics all share, in their semantics, the general schema of  $\exists x \Box \varphi$ , e.g., knowing how to achieve  $\varphi$  roughly means that there exists a way such that you know that it is a way to ensure that  $\varphi$ . Moreover, the resulting logics are decidable. Inspired by those particular logics, in this work, we propose a very general and powerful framework based on quantifier-free predicate language extended by a new modality  $\Box^x$ , which packs exactly  $\exists x \Box$  together. We show that the resulting language, though much more expressive, shares many good properties of the basic propositional modal logic over arbitrary models, such as finite-tree-model property and van Benthem-like characterization w.r.t. first-order modal logic. We axiomatize the logic over S5 frames with intuitive axioms to capture the interaction between  $\Box^x$  and know-that operator in an epistemic setting.

## 1 Introduction

Standard epistemic logic studies valid reasoning patterns about *knowing that*. However, in natural language, knowledge is also expressed by *knowing whether*, *knowing how*, *knowing what*, *knowing why* and so on. Recent years witnessed a growing interest in the non-standard epistemic logics of such expressions (cf. e.g., [32, 41, 42, 16, 7, 8, 38, 40, 29, 39, 11, 30, 28, 2] and the survey [39]).<sup>1</sup> In this line of work, various new modalities of know-wh are introduced,<sup>2</sup> all of which share the general *de re* schema  $\exists x \Box \varphi(x)$  in their semantics, e.g., “knowing how to achieve  $\varphi$ ” roughly means that there *exists* a way such that you *know that* it is a way to ensure that  $\varphi$  [40]; “knowing why  $\varphi$ ” means that there *exists* an explanation such that you *know that* it is an explanation to the fact  $\varphi$  [44]. Actually, in the early days of epistemic logic, Hintikka already used such formulations to handle *knowing who* [20] (cf. the survey [39] for a detailed discussion on Hintikka’s early contributions). Such interpretations are grounded also in philosophy and linguistics (cf. e.g., [36, 35]).

Such a semantic schema is based on the so-called *mention-some* interpretation to the wh-questions embedded in those knowledge expressions [17]. There is also a *mention-all* interpretation [15], which makes sense in many other situations, e.g., “knowing who came to the party” means, under an exhaustive reading, that for *each* relevant person, you know whether (s)he came to the party or not, which can be summarized as  $\forall x(\Box \varphi(x) \vee \Box \neg \varphi(x))$ . There are degenerated cases when the two interpretations coincide, e.g., “knowing [what] the value of  $c$  [is]” means, under the interpretation of mention-some, that there *exists* a value such that you *know that* it is the value of  $c$ , which is equivalent to the mention-all interpretation: for any value, you know whether it is the value of  $c$ , given there is *one and only one* real value of  $c$ .

---

\*The author thanks Pavel Naumov for insightful initial discussions on the language and semantics of this framework, and thanks Timothy Williamson for pointing out the categorization of know-wh by *mention-all* and *mention-some*. The author is also grateful to Yanjun Li and Jixin Liu for carefully reading an earlier draft of the paper.

<sup>1</sup>This is not meant to be an exhaustive list, e.g., see also [26, 19, 25] about logics of knowing how in the setting of ATL.

<sup>2</sup>Know-wh means verb *know* followed by a wh-word.

Given the experiences in dealing with those particular cases of know-wh, it is the time to lay out a general background framework for the shared “logical core” of those logics.<sup>3</sup> This paper is the initial step towards this purpose by extending the predicate language with the *mention-some* operator  $\Box^x$ , which is essentially a package of  $\exists x\Box$ . With the variable in place, we can say much more than those existing logics of know-wh, e.g., “I know a theorem of which I do not know any proof”  $\Box^x\neg\Box^y\textit{Prove}(y,x)$ , or, in a multi-agent setting, “*i* knows a country which *j* knows its capital”:  $\Box_i^x\Box_j^y\textit{Capital}(y,x)$ . Actually, when  $x$  does not appear in  $\varphi$ ,  $\Box^x\varphi$  is equivalent to  $\Box\varphi$ , thus our language is indeed an extension of the standard modal language. Moreover, we will show that in the epistemic setting, this mention-some operator can also express mention-all. By having the predicate symbols in the language, we can also talk about the *content* of knowledge, which may be useful to bridge epistemic logic and knowledge representation. We also believe that the new modality can be interpreted meaningfully not only in the epistemic context.

From a technical point of view, what we discovered is a well-behaved yet powerful fragment of first-order modal logic (over arbitrary models), as the following main technical contributions of this paper demonstrate:

- We propose a novel notion of bisimulation which can characterize the expressive power of our language within first-order modal logic precisely, over arbitrary models.
- Over arbitrary models, satisfiability of the equality-free fragment is not only decidable but also PSPACE-complete, just as the complexity of basic propositional modal logic, demonstrated by a tableau-like method to show some strong finite-tree-model property.
- We give a sound and complete proof system to our logic over epistemic (S5) models with the equivalence relation. However, we show that over S5 models, our logic is undecidable.

Due to historical and technical reasons, first-order modal logic (**FOML**), in particular in the epistemic setting, has not been thoroughly studied as its propositional brother (cf. e.g., [6, 12, 39]). Decidable fragments of first-order modal logic are usually obtained by restricting the occurrences of variables (particularly in the scope of  $\Box$ ) (cf. e.g., [22, 23, 21, 3]). We hope our framework and techniques can pave a new way to some interesting fragments of first-order modal logic, using new modalities to pack quantifiers and standard modalities together, which also reflects the “secret of success” of basic propositional modal logic as a nicely balanced logic between expressivity and complexity.

In the rest of this paper, we introduce the language and semantics of our framework in Section 2, study its expressivity over arbitrary models in Section 3, prove the complexity of the equality-free fragment in Section 4, give the proof systems in the epistemic setting in Section 5, and prove their completeness and undecidability in Section 6.

## 2 Syntax and Semantics

Throughout this paper we assume a fixed countably infinite set of variables  $\mathbf{X}$ , and a fixed set of predicate symbols  $\mathbf{P}$ . Furthermore, we assume that each predicate symbol is associated with a unique non-negative integer called its *arity*. We use  $\bar{x}$  to denote a finite sequence of (distinct) variables in  $\mathbf{X}$ , in the order of a fix enumeration of  $\mathbf{X}$ . By abusing the notation, we also view  $\bar{x}$  as a set of variables in  $\bar{x}$  occasionally. In this paper, for the brevity of presentation, we focus on the following unimodal language, but the results and techniques can be generalized to the polymodal language.<sup>4</sup>

<sup>3</sup>Note that this does not imply that we can simply use a single framework to cover all of those particular cases, since the details of the semantics in each setting matter a lot in deciding the characteristic axioms and rules in each different setting. Moreover, for example, in the setting of *knowing how* logics, we need to quantify over second-order objects (plans or strategies).

<sup>4</sup>We leave the discussion about the extension with constants or function symbols to the full version of the paper.

**Definition 2.1 (Language  $\text{MLMS}^{\approx}$ )** Given  $\mathbf{X}$  and  $\mathbf{P}$ ,

$$\varphi ::= x \approx y \mid P\bar{x} \mid \neg\varphi \mid (\varphi \wedge \psi) \mid \Box^x \varphi$$

where  $x, y \in \mathbf{X}$ ,  $P \in \mathbf{P}$ . We call the equality-free fragment  $\text{MLMS}$  (modal logic of mention-some).

$\Box^x \varphi$  can be read as *knowing some  $x$  such that  $\varphi(x)$*  in the epistemic context. We use the usual abbreviations  $\top, \perp, \vee, \rightarrow$ , and write  $\Diamond^x$  for  $\neg\Box^x\neg$ , i.e., the dual of  $\Box^x$ . We define the *free and bound occurrences* of variables as in first-order logic by viewing  $\Box^x$  as a quantifier. We call  $x$  a *free variable* in  $\varphi$  ( $x \in \text{FV}(\varphi)$ ), if there is a free occurrence of  $x$  in  $\varphi$ . We write  $\varphi(\bar{x})$  if all the free variables in  $\varphi$  are included in  $\bar{x}$ . Given an  $\text{MLMS}^{\approx}$  formula  $\varphi$  and  $x, y \in \mathbf{X}$ , we write  $\varphi[y/x]$  for the formula obtained by replacing every free occurrence of  $x$  by  $y$ . To simplify the discussion, we do not include constant symbols and function symbols in the language and leave them to a future occasion.<sup>5</sup>

As for the semantics, to be general enough, following [4, 45] we use the first-order Kripke model with an *increasing domain*, and flesh out the intuitive idea of mention-some discussed in the introduction.

**Definition 2.2** An (increasing domain) model  $\mathcal{M}$  for  $\text{MLMS}^{\approx}$  is a tuple  $\langle W, D, \delta, R, \rho \rangle$  where:

$W$  is a non-empty set.

$D$  is a non-empty set.

$R \subseteq W \times W$  is a binary relation over  $W$ .

$\delta : W \rightarrow 2^D$  assigns to each  $w \in W$  a non-empty local domain s.t.  $wRv$  implies  $\delta(w) \subseteq \delta(v)$  for any  $w, v \in W$ . We also write  $D_w$  for  $\delta(w)$ .

$\rho : \mathbf{P} \times W \rightarrow \bigcup_{n \in \omega} 2^{D^n}$  such that  $\rho$  assigns each  $n$ -ary predicate on each world an  $n$ -ary relation on  $D$ .<sup>6</sup>

A *constant domain model* is a model such that  $D_w = D$  for any  $w \in W$ . A *finite model* is a model with both a finite  $W$  and a finite  $D$ . Given a model  $\mathcal{M}$ , we denote its components as  $W^{\mathcal{M}}, D^{\mathcal{M}}, \delta^{\mathcal{M}}, R^{\mathcal{M}}$ , and  $\rho^{\mathcal{M}}$ . To interpret free variables, we also need a variable assignment  $\sigma : \mathbf{X} \rightarrow D$ . The formulas are interpreted on models with variable assignments.

$\mathcal{M}, w, \sigma \models x \approx y$	$\Leftrightarrow$	$\sigma(x) = \sigma(y)$
$\mathcal{M}, w, \sigma \models P(x_1 \cdots x_n)$	$\Leftrightarrow$	$(\sigma(x_1), \dots, \sigma(x_n)) \in \rho(P, w)$
$\mathcal{M}, w, \sigma \models \neg\varphi$	$\Leftrightarrow$	$\mathcal{M}, w, \sigma \not\models \varphi$
$\mathcal{M}, w, \sigma \models (\varphi \wedge \psi)$	$\Leftrightarrow$	$\mathcal{M}, w, \sigma \models \varphi$ and $\mathcal{M}, w, \sigma \models \psi$
$\mathcal{M}, w, \sigma \models \Box^x \varphi$	$\Leftrightarrow$	there exists an $a \in \delta(w)$ such that $\mathcal{M}, v, \sigma[x \mapsto a] \models \varphi$ for all $v$ s.t. $wRv$

where  $\sigma[x \mapsto a]$  denotes another assignment just like  $\sigma$  except mapping  $x$  to  $a$ .

As an intuitive definition, following [24], we say  $\varphi$  is *valid*, if  $\varphi$  is true on any  $\mathcal{M}, w$  w.r.t. any  $\sigma$  such that  $\sigma(x) \in \delta^{\mathcal{M}}(w)$  for all  $x \in \mathbf{X}$ . Correspondingly,  $\varphi$  is *satisfiable* if  $\neg\varphi$  is not valid, i.e.,  $\mathcal{M}, w, \sigma \models \varphi$  for some  $\mathcal{M}, w$  and  $\sigma$  such that  $\sigma(x) \in \delta^{\mathcal{M}}(w)$  for all  $x \in \mathbf{X}$ .

<sup>5</sup>Note that constant and functions can be coded using  $\mathbf{P}$  and  $\approx$  in full  $\text{FOML}^{\approx}$  (cf. e.g., [24]). However, our language is a fragment of  $\text{FOML}^{\approx}$ .

<sup>6</sup>Following [24, 6], we do not require that the interpretation of  $P$  at a world is based on the local domain. Actually, as we will see later in Corollary 4.1, this seemingly ‘counterintuitive’ generalization does not affect the satisfiability or validity: each satisfiable formula  $\varphi$  is satisfiable in a model where  $\rho(P, w)$  is based on objects in  $D_w$  (cf. also [24]).

It is not hard to see that if  $\sigma(x) = \sigma'(x)$  for all the free  $x$  in  $\varphi$ ,  $\mathcal{M}, w, \sigma \models \varphi \Leftrightarrow \mathcal{M}, w, \sigma' \models \varphi$ . In this light, we write  $\mathcal{M}, w \models \varphi[\bar{a}]$  to denote  $\mathcal{M}, w, \sigma \models \varphi(\bar{x})$  for any  $\sigma$  such that  $\sigma$  assigns free variables  $\bar{x}$  of  $\varphi$  the corresponding objects in  $\bar{a}$  given  $|\bar{x}| = |\bar{a}|$ , where  $|\cdot|$  denotes the length.

For comparison, the standard semantics for  $\Box$  is defined as:

$$\boxed{\mathcal{M}, w, \sigma \models \Box\varphi \Leftrightarrow \text{for all } v \text{ such that } wRv \mathcal{M}, v, \sigma \models \varphi}$$

Truth conditions of  $\Box^x\varphi$  and  $\Diamond^x\varphi$  can then be defined using  $\Box, \Diamond$ :

$$\boxed{\begin{array}{l} \mathcal{M}, w, \sigma \models \Box^x\varphi \Leftrightarrow \text{there exists an } a \in \delta(w) \text{ such that } \mathcal{M}, w, \sigma[x \mapsto a] \models \Box\varphi \\ \mathcal{M}, w, \sigma \models \Diamond^x\varphi \Leftrightarrow \text{for all } a \in \delta(w), \mathcal{M}, w, \sigma[x \mapsto a] \models \Diamond\varphi \end{array}}$$

It is now clear that  $\Box\varphi$  is equivalent to  $\Box^x\varphi$  where  $x \notin FV(\varphi)$ . Therefore  $\mathbf{MLMS}^\approx$  can be viewed as an extension of the basic propositional modal language. Therefore, in the context of  $\mathbf{MLMS}^\approx$ ,  $\Box\varphi$  can be viewed as an abbreviation. It also becomes evident that  $\Box^x\varphi$  and  $\Diamond^x\varphi$  are essentially  $\exists x\Box\varphi$  and  $\forall x\Diamond\varphi$  in  $\mathbf{FOML}^\approx$  respectively. In the following we study the expressivity of  $\mathbf{MLMS}^\approx$  in comparison with  $\mathbf{FOML}^\approx$ .

### 3 Expressivity

Note that our semantics for  $\Box^x$  has the  $\exists\forall$  pattern which is similar to the neighbourhood semantics for modal logic [24]. Moreover,  $\mathbf{MLMS}^\approx$  can be viewed as a fragment of the corresponding  $\mathbf{FOML}^\approx$ . Indeed, inspired by the *world-object bisimulation* for  $\mathbf{FOML}^\approx$  [4, 45] and bisimulation for monotonic neighbourhood modal logic [31, 18], we propose a novel notion of bisimulation for  $\mathbf{MLMS}^\approx$ . Before the formal definition, given a model  $\mathcal{M}$ , let  $D_{\mathcal{M}}^*$  be the set of (possibly empty) finite sequence of objects in  $D^{\mathcal{M}}$ . A *partial isomorphism* between  $\bar{a} \in D_{\mathcal{M}}^*$  and  $\bar{b} \in D_{\mathcal{N}}^*$  such that  $|a| = |b|$  is an isomorphism mapping  $a_i$  to  $b_i$  w.r.t. relevant interpretations of predicates (cf. e.g, [4]). It is *partial* since it is not about all the objects in  $D^{\mathcal{M}}$  and  $D^{\mathcal{N}}$ .

**Definition 3.1 ( $\exists\Box$ -Bisimulation)** *Given two models  $\mathcal{M}$  and  $\mathcal{N}$ , the relation  $Z \subseteq (W^{\mathcal{M}} \times D_{\mathcal{M}}^*) \times (W^{\mathcal{N}} \times D_{\mathcal{N}}^*)$  is call an  $\exists\Box$ -bisimulation, if for every  $((w, \bar{a}), (v, \bar{b})) \in Z$  such that  $|\bar{a}| = |\bar{b}|$  the following holds (for brevity, comma in  $(w, \bar{a})$  is omitted):*

*PISO*  $\bar{a}$  and  $\bar{b}$  form a partial isomorphism w.r.t. identity and interpretations of predicates at  $w$  and  $v$  respectively.

$\exists\Box\text{Zig}$  For any  $c \in D_w^{\mathcal{M}}$ , there is a  $d \in D_v^{\mathcal{N}}$  such that for any  $v' \in W^{\mathcal{N}}$  if  $vRv'$  then there exists  $w' \in W^{\mathcal{M}}$  such that  $wRw'$  and  $w'\bar{a}cZv'\bar{b}d$ .

$\exists\Box\text{Zag}$  For any  $d \in D_v^{\mathcal{N}}$ , there is a  $c \in D_w^{\mathcal{M}}$  such that for any  $w' \in W^{\mathcal{M}}$  if  $wRw'$  then there exists  $v' \in W^{\mathcal{N}}$  such that  $vRv'$  and  $w'\bar{a}cZv'\bar{b}d$ .

We say  $\mathcal{M}, \bar{w}\bar{a}$  and  $\mathcal{N}, \bar{v}\bar{b}$  are  $\exists\Box$ -bisimilar ( $\mathcal{M}, \bar{w}\bar{a} \leftrightarrow_{\exists\Box} \mathcal{N}, \bar{v}\bar{b}$ ) if  $|a| = |b|$  and there is an  $\exists\Box$ -bisimulation linking  $\bar{w}\bar{a}$  and  $\bar{v}\bar{b}$ . In particular, we say  $\mathcal{M}, w$  and  $\mathcal{N}, v$  are  $\exists\Box$ -bisimilar if  $\mathcal{M}, w \leftrightarrow_{\exists\Box} \mathcal{N}, v$ , i.e., when  $|\bar{a}| = |\bar{b}| = 0$ .

It is not hard to show that  $\leftrightarrow_{\exists\Box}$  is indeed an equivalence relation.

Note that our bisimulation notion is much weaker than isomorphism, in particular the domains of the two bisimilar models do not necessarily have the same cardinality.

**Example 1** Consider the constant domain models  $\mathcal{M}, \mathcal{N}$ :

$$\mathcal{M} : \quad \begin{array}{c} \underline{w} \longrightarrow v : Pa \\ \searrow \\ u : Pb \end{array} \quad \mathcal{N} : \quad \begin{array}{c} \underline{s} \longrightarrow t : Pc \\ \searrow \\ r \end{array}$$

where  $D^{\mathcal{M}} = \{a, b\}$ ,  $D^{\mathcal{N}} = \{c\}$ ,  $\rho^{\mathcal{M}}(P, w) = \rho^{\mathcal{N}}(P, s) = \rho^{\mathcal{N}}(P, r) = \emptyset$ ,  $\rho^{\mathcal{M}}(P, v) = \{a\}$ ,  $\rho^{\mathcal{M}}(P, u) = \{b\}$ ,  $\rho^{\mathcal{N}}(P, t) = \{c\}$ . Suppose  $P$  is the only predicate, we can show that  $\mathcal{M}, w \not\leftrightarrow_{\exists\Box} \mathcal{N}, s$  by an  $\exists\Box$ -bisimulation  $Z$  (pay attention to the switch of the two models in the second half of the definitions of  $\exists\Box\text{Zig}$  and  $\exists\Box\text{Zag}$ ):

$$\{(w, s), (va, tc), (ub, tc), (vb, rc), (ua, rc)\}$$

Also note that  $\exists\Box\text{Zig}$  and  $\exists\Box\text{Zag}$  hold trivially for  $w\bar{a}$  and  $v\bar{b}$  if  $w$  and  $v$  do not have any successor, based on the fact that local domains are non-empty by definition.

We write  $\mathcal{M}, w\bar{a} \equiv_{\text{MLMS}^\approx} \mathcal{N}, v\bar{b}$  if  $|\bar{a}| = |\bar{b}|$  and for all any  $\text{MLMS}^\approx$  formula  $\varphi(\bar{x})$  such that  $|\bar{x}| = |\bar{a}|$ :  $\mathcal{M}, w \models \varphi[\bar{a}] \Leftrightarrow \mathcal{N}, v \models \varphi[\bar{b}]$ . We can show that  $\text{MLMS}^\approx$  is invariant under  $\exists\Box$ -bisimilarity.

**Theorem 3.1**  $\mathcal{M}, w\bar{a} \leftrightarrow_{\exists\Box} \mathcal{N}, v\bar{b}$  then  $\mathcal{M}, w\bar{a} \equiv_{\text{MLMS}^\approx} \mathcal{N}, v\bar{b}$ .

**Proof:** It suffices to prove that if  $Z$  is an  $\exists\Box$ -bisimulation linking  $w\bar{a}$  and  $v\bar{b}$  such that  $|\bar{a}| = |\bar{b}|$ , then  $\mathcal{M}, w \models \varphi[\bar{a}] \Leftrightarrow \mathcal{N}, v \models \varphi[\bar{b}]$ . Given a bisimulation  $Z$ , we do induction on the structure of  $\varphi$ . Supposing  $\varphi$  is an atomic formula, due to PISO,  $\mathcal{M}, w \models \varphi[\bar{a}]$  iff  $\mathcal{N}, v \models \varphi[\bar{b}]$  for any atomic formula  $\varphi(\bar{x})$ .

Suppose  $\varphi$  is in the shape of  $\Box^x \psi(\bar{x})$  and  $\mathcal{M}, w \models \varphi[\bar{a}]$ . By the semantics there is a  $c \in D_w^{\mathcal{M}}$  such that  $\mathcal{M}, w' \models \psi[\bar{a}c]$  for all the  $w'$  such that  $wRw'$ . According to  $\exists\Box\text{Zig}$ , there exists a  $d \in D_v^{\mathcal{N}}$  such that the second half of the condition holds. We claim that  $\mathcal{N}, v' \models \psi[\bar{b}d]$  for any  $v'$  such that  $vRv'$ . Suppose not, then there is a  $v'$  such that  $\mathcal{N}, v' \not\models \psi[\bar{b}d]$  and  $vRv'$ . Then according to  $\exists\Box\text{Zig}$ , there is  $w'$  such that  $wRw'$  and  $w'\bar{a}cZv'\bar{b}d$ . By IH,  $\mathcal{M}, w' \not\models \psi[\bar{a}c]$ . Contradiction.  $\blacksquare$

**Corollary 3.1**  $\mathcal{M}, w \leftrightarrow_{\exists\Box} \mathcal{N}, v$  then for any closed  $\text{MLMS}^\approx$ -formula  $\varphi$ :  $\mathcal{M}, w \models \varphi \Leftrightarrow \mathcal{N}, v \models \varphi$ .

By the above invariance results, we can show that many natural combinations of quantifiers and modalities are not expressible in  $\text{MLMS}^\approx$ .

**Proposition 3.1**  $\Box\exists Px$ ,  $\exists x\Diamond Px$  and  $\Diamond\exists Px$  are not expressible in  $\text{MLMS}^\approx$ .

**Proof:** In this proof, we again consider constant domain models. For  $\Box\exists Px$ , consider the bisimilar models in Example 1.  $\Box\exists Px$  holds on  $\mathcal{M}, w$  but not on  $\mathcal{N}, s$ , thus it is not expressible in  $\text{MLMS}^\approx$ . For  $\exists x\Diamond Px$  and  $\Diamond\exists Px$ , consider:

$$\mathcal{M} : \quad \begin{array}{c} \underline{w} \longrightarrow v : Pa \\ \searrow \\ u \end{array} \quad \mathcal{N} : \quad \underline{s} \longrightarrow t$$

where  $D^{\mathcal{M}} = \{a, b\}$ ,  $D^{\mathcal{N}} = \{c\}$  as before,  $\rho^{\mathcal{M}}(P, w) = \rho^{\mathcal{M}}(P, u) = \rho^{\mathcal{N}}(P, t) = \rho^{\mathcal{N}}(P, s) = \emptyset$ ,  $\rho^{\mathcal{M}}(P, v) = \{a\}$ . Clearly,  $\exists x\Diamond Px$  and  $\Diamond\exists Px$  are true at  $\mathcal{M}, w$  but false at  $\mathcal{N}, s$ . However, we can show that  $\mathcal{M}, w \leftrightarrow_{\exists\Box} \mathcal{N}, s$  by an  $\exists\Box$ -bisimulation  $Z$ :

$$\{(w, s), (ua, tc), (vb, tc), (ub, tc)\}.$$

To precisely characterize  $\text{MLMS}^\approx$  within the corresponding  $\text{FOML}^\approx$ , we need a notion of saturation. In the following, we write  $\Gamma(\bar{x})$  if all the free variables in the set of  $\text{MLMS}^\approx$ -formulas  $\Gamma$  are included in  $\bar{x}$ . Inspired by [45], we can generalize the concept of m-saturation for propositional modal logic (cf. [5]) as follows:

**Definition 3.2** A model  $\mathcal{M}$  is said to be  $\exists\Box$ -saturated, if for any  $w \in W^{\mathcal{M}}$ , and any finite sequence  $\bar{a} \in D_{\mathcal{M}}^*$ , the following two conditions are satisfied:

- $\exists\Box$ -type If for each finite subset  $\Delta$  of a set  $\Gamma(\bar{y}x)$  where  $|\bar{y}| = |\bar{a}|$ ,  $\mathcal{M}, w \models \Box^x \bigwedge \Delta[\bar{a}]$ ,<sup>7</sup> then there is a  $c \in D_w^{\mathcal{M}}$  such that  $\mathcal{M}, w \models \Box\varphi[\bar{a}c]$  for all  $\varphi \in \Gamma$ , where  $x$  is assigned  $c$ .
- $\Diamond$ -type If for each finite subset  $\Delta$  of  $\Gamma(\bar{x})$  such that  $|\bar{x}| = |\bar{a}|$ ,  $\mathcal{M}, w \models \Diamond \bigwedge \Delta[\bar{a}]$ , then  $\mathcal{M}, w \models \Diamond\varphi[\bar{a}]$  for each  $\varphi \in \Gamma$ .

Note that in the above definition, to simplify the presentation, we use  $\Box, \Diamond$  which are expressible in  $\text{MLMS}^{\approx}$ .  $\Diamond$ -type condition is essentially the m-saturation adapted with variable assignments. Recall that a finite model has a finite domain and finitely many worlds. It can be verified that:<sup>8</sup>

**Proposition 3.2** Every finite model is  $\exists\Box$ -saturated.

We can obtain the Hennessy-Milner-type theorem (cf. [5]) to establish the equality between  $\exists\Box$ -bisimilarity and  $\text{MLMS}^{\approx}$ -equivalence.

**Theorem 3.2** For  $\exists\Box$ -saturated models  $\mathcal{M}, \mathcal{N}$  and  $|\bar{a}| = |\bar{b}|$ :

$$\mathcal{M}, w\bar{a} \xleftrightarrow{\exists\Box} \mathcal{N}, v\bar{b} \Leftrightarrow \mathcal{M}, w\bar{a} \equiv_{\text{MLMS}^{\approx}} \mathcal{N}, v\bar{b}$$

**Proof:** Due to Theorem 3.1, we only need to show the right-to-left direction. We define  $Z = \{(w\bar{a}, v\bar{b}) \mid w \in W^{\mathcal{M}}, v \in W^{\mathcal{N}}, \bar{a} \in D_{\mathcal{M}}^*, \bar{b} \in D_{\mathcal{N}}^*, |\bar{a}| = |\bar{b}|, \mathcal{M}, w\bar{a} \equiv_{\text{MLMS}^{\approx}} \mathcal{N}, v\bar{b}\}$ . We need to show that  $Z$  is an  $\exists\Box$ -bisimulation. PISO is straightforward as  $\bar{a}$  and  $\bar{b}$  are finite and partial isomorphism between them can be expressed by atomic formulas. We only show  $\exists\Box\text{Zig}$  since  $\exists\Box\text{Zag}$  is similar. Assuming  $w\bar{a}Zv\bar{b}$  and there is a  $c \in D_w^{\mathcal{M}}$ . Let  $\Gamma = \{\varphi(\bar{y}x) \mid w \models \Box\varphi[\bar{a}c], |\bar{y}| = |\bar{a}|\}$ . Now for any finite set  $\Delta \subseteq \Gamma$  we have  $w \models \Box^x \bigwedge \Delta[\bar{a}]$ . Since  $w\bar{a} \equiv_{\text{MLMS}^{\approx}} v\bar{b}$ ,  $v \models \Box^x \bigwedge \Delta[\bar{b}]$ . Now by  $\exists\Box$ -type condition, we know there is a  $d \in D_v^{\mathcal{N}}$  such that  $v \models \Box\varphi[\bar{b}d]$  for all  $\varphi \in \Gamma$  ( $\star$ ). Now take an arbitrary  $v'$  such that  $vRv'$ , we show there is a  $w'$  such that  $wRw'$  and  $w'\bar{a}c \equiv_{\text{MLMS}^{\approx}} v'\bar{b}d$ . Let  $\Sigma = \{\varphi(\bar{x}) \mid \mathcal{N}, v' \models \varphi[\bar{b}d], |\bar{x}| = |\bar{b}d|\}$ . It is clear that for each finite set  $\Delta \subseteq \Sigma$ ,  $v' \models \Diamond \bigwedge \Delta[\bar{b}d]$ , namely  $v' \models \neg\Box\neg \bigwedge \Delta[\bar{b}d]$ . We can show  $w \models \neg\Box\neg \bigwedge \Delta[\bar{a}c]$ . Suppose not, then  $w \models \Box\neg \bigwedge \Delta[\bar{a}c]$ , thus  $\neg \bigwedge \Delta \in \Gamma$ . By ( $\star$ ),  $v \models \Box\neg \bigwedge \Delta[\bar{b}d]$ , which is in contradiction with  $v \models \neg\Box\neg \bigwedge \Delta[\bar{b}d]$ . Now it is clear that  $w \models \Diamond \bigwedge \Delta[\bar{a}c]$  for each finite  $\Delta \subseteq \Sigma$ . By the  $\Diamond$ -type condition of  $\exists\Box$ -saturation, there is a successor  $w'$  of  $w$  such that  $w' \models \Sigma[\bar{a}c]$ . Therefore  $w'\bar{a}c \equiv_{\text{MLMS}^{\approx}} v'\bar{b}d$ , namely  $(w'\bar{a}c, v'\bar{b}d) \in Z$ . This completes the proof for  $\exists\Box\text{Zig}$ . ■

**Corollary 3.2** For finite models:  $\mathcal{M}, w\bar{a} \xleftrightarrow{\exists\Box} \mathcal{N}, v\bar{b} \Leftrightarrow \mathcal{M}, w\bar{a} \equiv_{\text{MLMS}^{\approx}} \mathcal{N}, v\bar{b}$ . For  $\exists\Box$ -saturated models:  $\mathcal{M}, w \xleftrightarrow{\exists\Box} \mathcal{N}, v \Leftrightarrow \mathcal{M}, w$  and  $\mathcal{N}, v$  satisfies the same closed  $\text{MLMS}^{\approx}$ -formulas (sentences).

Now we can characterize  $\text{MLMS}^{\approx}$  in the corresponding first-order language.

**Definition 3.3** Given  $X, P$  as before, the corresponding  $\text{FOML}^{\approx}$  language is defined as follows:

$$\varphi ::= x \approx y \mid P\bar{x} \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \forall x\varphi \mid \Box\varphi$$

The corresponding two-sorted first-order language  $2\text{SFOL}^{\approx}$  is:

$$\varphi ::= x \approx y \mid Q\bar{u}\bar{x} \mid R\bar{u}\bar{v} \mid E\bar{u}\bar{x} \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \forall x\varphi \mid \forall u\varphi$$

where  $x, y \in X$  and  $u, v \in Y$  which is a collection of world variables disjoint from  $X$ ;  $Q \in \mathbf{Q}$  and  $\mathbf{Q}$  is the smallest collection of predicate symbols such that for each  $n$ -ary  $P \in \mathbf{P}$  there is a unique  $Q_P \in \mathbf{Q}$  with the arity  $n + 1$ .  $E$  is a new predicate symbol to say which object exists on which world.

<sup>7</sup>Here we require  $\bar{y}$  are assigned  $\bar{a}$  correspondingly. This is only to avoid introducing extended language with constants for  $\bar{a}$  as in standard model theory, since we did not define interpretation for constant symbols. Similarly below.

<sup>8</sup>Due to limited space, we omit some proofs and leave them to the extended version of this conference paper.

It is trivial to define a translation  $r$  from  $\mathbf{MLMS}^{\approx}$  to the corresponding  $\mathbf{FOML}^{\approx}$  by setting  $r(\Box^x \varphi) = \exists x \Box r(\varphi)$ . It is clear that this translation is truth preserving. To translate  $\mathbf{MLMS}^{\approx}$  into the formulas in the corresponding  $\mathbf{2SFOL}^{\approx}$  (with  $u, v$  as the only world variables), we define the following  $t_u$  inspired by [4]:

**Definition 3.4 (Translation from  $\mathbf{MLMS}^{\approx}$  to  $\mathbf{2SFOL}^{\approx}$ )**

$$\begin{aligned} t_u(x \approx y) &= x \approx y & t_u(P\bar{x}) &= Q_P(u, \bar{x}) \\ t_u(\neg \psi) &= \neg t_u(\psi) & t_u(\varphi \wedge \psi) &= t_u(\varphi) \wedge t_u(\psi). \\ t_u(\Box^x \psi) &= \exists x (Eux \wedge \forall v (Ruv \rightarrow t_v(\psi))) \end{aligned}$$

$t_v$  is defined symmetrically, by swapping  $u$  and  $v$ .

In this way, every  $\mathbf{MLMS}^{\approx}$ -formula is translated into a two-world-variable formula of  $\mathbf{2SFOL}^{\approx}$  with one free variable, similar to the standard translation of basic modal language to first-order language.

We can also view a model for  $\mathbf{MLMS}^{\approx}$  as a model for  $\mathbf{2SFOL}^{\approx}$  by turning  $\rho(w, P)$  into the interpretation of the corresponding  $Q_P$  in the most natural way. It is not hard to show:

**Proposition 3.3** For any  $\mathbf{MLMS}^{\approx}$  formula  $\varphi$ :

$$\mathcal{M}, w, \sigma \models \varphi \Leftrightarrow \mathcal{M}, \sigma' \Vdash t_u(\varphi)$$

where  $\sigma'$  is  $\sigma$  extended by  $u \mapsto w$ ,  $\Vdash$  is the standard semantics for  $\mathbf{2SFOL}^{\approx}$  (cf. e.g. [6]).

It follows from the compactness of  $\mathbf{2SFOL}^{\approx}$  that:<sup>9</sup>

**Proposition 3.4**  $\mathbf{MLMS}^{\approx}$  is compact.

Based on Theorem 3.2 and the compactness of  $\mathbf{MLMS}^{\approx}$ , it is relatively routine to prove the van-Benthem-like characterization theorem using the strategy in [4, 5], which makes use of  $\omega$ -saturation in first-order model theory. We omit the proof due to limited space.

**Theorem 3.3** A  $\mathbf{2SFOL}^{\approx}$ -formula  $\varphi(\bar{x}u)$  is equivalent to an  $\mathbf{MLMS}^{\approx}$ -formula (over  $\mathbf{MLMS}^{\approx}$ -models) iff it is invariant under  $\exists \Box$ -bisimilarity.

Since any  $\mathbf{FOML}^{\approx}$  formula can also be viewed as a  $\mathbf{2SFOL}^{\approx}$  formula  $\varphi(\bar{x}u)$ , via a natural translation like  $t_u$  (cf. [6, 45]), it follows that:

**Corollary 3.3** A  $\mathbf{FOML}^{\approx}$ -formula is equivalent to an  $\mathbf{MLMS}^{\approx}$ -formula iff it is invariant under  $\exists \Box$ -bisimilarity.

## 4 Satisfiability of $\mathbf{MLMS}$

Note that by  $t_u$  we can translate an  $\mathbf{MLMS}^{\approx}$  formula to an equivalent  $\mathbf{2SFOL}^{\approx}$  formula, and eventually to a  $\mathbf{FOL}^{\approx}$  formula with one free world variable  $u$  (see Footnote 9). For example  $\Box^x Px$  becomes

$$\varphi(u) = \exists x (S_1x \wedge S_2u \wedge Eux \wedge \forall v ((S_2v \wedge Ruv) \rightarrow Q_P vx))$$

<sup>9</sup> $\mathbf{2SFOL}^{\approx}$  inherits the compactness from first-order logic since there is a truth preserving translation from  $\mathbf{2SFOL}^{\approx}$  to  $\mathbf{FOL}^{\approx}$  by using two new unary predicates  $S_1, S_2$  for the two sorts, together with a single unrestricted quantifiers to mimic the two-sorted quantifiers. Note that we can also express the constraints on the sorts in  $\mathbf{FOL}^{\approx}$  by  $\theta(S_1, S_2) = \forall x ((S_1x \vee S_2x) \wedge \neg(S_1x \wedge S_2x))$ .

in the corresponding  $\mathbf{FOL}^{\approx}$  language. Note that since we only consider increasing domain models,  $\varphi$  is equivalent to

$$\varphi(u) = \exists x(S_1x \wedge S_2u \wedge Eux \wedge \forall v((S_2v \wedge Ruv \wedge Evx) \rightarrow Q_Pvx))$$

which is in the (loosely) guarded fragment of first-order logic known to be decidable [1]. However, it does not directly imply that our logic is decidable, since, for example, to capture the increasing domain model we do need the following first-order constraint:

$$\chi = \forall u \forall v \forall x((S_2u \wedge S_2v \wedge S_1x \wedge Eux \wedge Ruv) \rightarrow Evx)$$

$\chi$  can be viewed as some form of transitivity, and it is not in the guarded fragment, since  $v$  and  $x$  are free in the consequent but they do not co-exist in any of the atomic guards in the antecedent. On the other hand, such a transitivity-like constrained predicate  $E$  only appears in the guards of the translations of  $\mathbf{MLMS}^{\approx}$ -formulas, which may be handled by the techniques in [37] for the guarded fragment with transitive guards known to be decidable. We leave the detailed discussion connecting  $\mathbf{MLMS}^{\approx}$  to guarded fragments to a future occasion.

In this section, we give an intuitive tableau-like method to decide whether an equality-free  $\mathbf{MLMS}$ -formula is satisfiable, inspired by the satisfiability games for various temporal logics [13, 27].<sup>10</sup>

For the ease of presentation, we consider the following equivalent language of  $\mathbf{MLMS}$  in *positive normal form (PNF)*, where negations only appear with atomic formulas:<sup>11</sup>

$$\varphi ::= P\bar{x} \mid \neg P\bar{x} \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid \Box^x \varphi \mid \Diamond^x \varphi$$

Note that for an arbitrary  $\mathbf{MLMS}$ -formula  $\varphi$ , we can rewrite it into an equivalent formula in PNF by the following rewriting rules and the replacement of equals:

$$r(\neg(\varphi \wedge \psi)) = r(\neg\varphi) \vee r(\neg\psi) \quad r(\neg\Box^x \varphi) = \Diamond^x r(\neg\varphi) \quad r(\neg\neg\varphi) = r(\varphi)$$

Note that  $|r(\varphi)| \leq 2|\varphi|$ .

In the following, we will focus on the formulas that are *clean* in the sense that no variable occurs both free and bound,<sup>12</sup> and no two distinct occurrences of modalities bind the same variable. As in first-order logic, it is easy to show that any  $\mathbf{MLMS}$  formula can be relettered into an equivalent clean formula with the same length, by renaming the bounded variables. We define the following tableau rules for all  $\mathbf{MLMS}$  formulas in PNF.

**Definition 4.1** *Tableau rules*

$\frac{w : \varphi_1 \vee \varphi_2, \Gamma, \sigma}{w : \varphi_1, \Gamma, \sigma \mid w : \varphi_2, \Gamma, \sigma} (\vee) \qquad \frac{w : \varphi_1 \wedge \varphi_2, \Gamma, \sigma}{w : \varphi_1, \varphi_2, \Gamma, \sigma} (\wedge)$
$\frac{\text{Given } n \geq 0, m \geq 1: \\ w : \Box^{x_1} \varphi_1, \dots, \Box^{x_n} \varphi_n, \Diamond^{y_1} \psi_1, \dots, \Diamond^{y_m} \psi_m, l_1 \dots l_k, \sigma}{\{(wv_{y_i}^x : \{\varphi_j \mid 1 \leq j \leq n\}, \psi_i[y/y_i], \sigma') \mid y \in \text{Dom}(\sigma'), i \in [1, m]\}} \text{ (BR)}$
$\frac{\text{Given } n \geq 1, k \geq 0: \\ w : \Box^{x_1} \varphi_1, \dots, \Box^{x_n} \varphi_n, l_1 \dots l_k, \sigma}{w : l_1 \dots l_k, \sigma} \text{ (END)}$

where  $\sigma' = \sigma \cup \{(x_j, x_j) \mid j \in [1, n]\}$  and  $l_k \in \text{lit}$  (the literals).

<sup>10</sup>See [33, 10] for tableau methods for first-order modal logic in general.

<sup>11</sup>PNF is often used in automata-theoretical methods to satisfiability problems of logics in computer science (cf. e.g., [34, 27]).

<sup>12</sup>In particular, if  $\Box^x$  or  $\Diamond^x$  appears in the formula then  $x$  does not have any free occurrence, even when  $\Box^x$  and  $\Diamond^x$  do not bind any occurrence of  $x$ , e.g.,  $(\Box^x P y) \wedge P x$  is not clean.



Those rules are used to generate tree-like structure, where the nodes are triples  $(w, \Gamma, \sigma)$  in which  $w$  is some name to denote a world,  $\Gamma$  is a finite set of **MLMS**-formulas, and  $\sigma$  is a partial function from  $\mathbf{X}$  to  $\mathbf{X}$  as an assignment. Note that Rule  $(\vee)$  is essentially a *choice*: given the numerator, you can only select *one* of the denominator to continue. On the other hand, the rule (BR) is a *branching* one, which generates all the nodes in the denominator set (even when  $n = 0$  i.e., the  $\Box^x$ -part is empty). It is a generalized version of the corresponding rule for basic modal logic (cf. [27]):

$$\frac{w : \Box\varphi, \Diamond\psi, l_1 \dots l_l}{v : \varphi, \psi}$$

The idea is that if there is a diamond formula then we need to generate a successor while keeping the information given by  $\Box$ -formulas. Here the complication is to manage the variable assignment properly. Note that if there are merely  $\Box^x$ -formulas without any  $\Diamond^x$ -formula, then we do not need to generate any successor, as captured by the rule END.

A *tableau* starting from  $(w : \Gamma, \sigma)$  is a tree where the successors of a node are generated by applying the rules, until no rule is applicable. Recall that you can only select one of the denominators to continue when applying  $(\vee)$ . Below is an example of a tableau where the (partial) function  $\sigma$  is represented by a set of ordered pairs:

$$\begin{array}{c} w : \{\Box^x(Px \vee Qx) \wedge \Diamond^y \neg Qy \wedge \neg Pz, \{(z, z)\}\} \quad (\wedge) \times 2 \\ | \\ w : \{\Box^x(Px \vee Qx), \Diamond^y \neg Qy, \neg Pz\} \{(z, z)\} \quad (\text{BR}) \\ / \quad \backslash \\ wv_y^x : \{Px \vee Qx, \neg Qx\}, \{(x, x), (z, z)\} \quad (\vee) \quad wv_y^z : \{Px \vee Qx, \neg Qz\}, \{(x, x), (z, z)\} \\ | \quad \quad \quad | \\ wv_y^x : \{Px, \neg Qx\}, \{(x, x), (z, z)\} \quad \quad wv_y^z : \{Qx, \neg Qz\}, \{(x, x), (z, z)\} \end{array}$$

Intuitively,  $(\vee)$  and  $(\wedge)$  will ‘decompose’ the formula until BR is applicable. It is not hard to see that all the leaf nodes are in the shape of  $(w : l_1 \dots l_k, \sigma)$ . Moreover, any tableau starting from  $(w : \Gamma, \sigma)$ , where  $\Gamma$  and  $\sigma$  are finite, is a finite tree, both in depth and width, since we only generate simpler formulas by the rules and the domain of the assignment is always finite. A tableau is called *open* if all its branches do not contain contradictions of literals at the same world, i.e. no  $P\bar{x}$  and  $\neg P\bar{x}$  appear together.

Given a tableau  $\mathcal{T}$ , we say a node  $(w : \Gamma, \sigma)$  is a *branching node* if it is branching due to the application of BR. We call  $(w : \Gamma, \sigma)$  the *last node of  $w$* , if it is a leaf node or a branching node. Clearly, given a  $w$  appearing in a tableau  $\mathcal{T}$ , the last node of  $w$  always uniquely exists, since we always only select one of the denominators for the rule  $(\vee)$ . We denote the last node of  $w$  in a given  $\mathcal{T}$  as  $t_w$ . Let

$$Dom(t_w) = \begin{cases} Dom(\sigma') & \text{if } t_w \text{ is branching} \\ Dom(\sigma) & \text{otherwise} \end{cases}$$

where  $\sigma$  is the assignment in  $t_w$  and  $\sigma'$  is the assignment defined as in BR w.r.t.  $t_w$ .

Actually, the open tableaux are pseudo models.

**Theorem 4.1** *For any clean MLMS-formula  $\theta$  in PNF, the following are equivalent:*

- *There is an open tableau from*

$$(r : \{\theta\}, \sigma_r = \{(x, x) \mid x \text{ is free in } \theta\} \cup \{(z, z)\})$$

where  $z \in X$  and it does not appear in  $\theta$ .

- $\theta$  is satisfiable in an increasing domain model.

**Proof:** Before the main proof, we need the following handy observations about any tableau  $\mathcal{T}$  starting from  $(r, \{\theta\}, \sigma_r)$  where  $\theta$  is clean. For any node  $(v : \Gamma, \sigma)$  in  $\mathcal{T}$ , we claim:

- (1) If  $\Box^x (\Diamond^x)$  occurs in  $\Gamma$ , then it only occurs once and there is no  $\Diamond^x (\Box^x)$  occurring in  $\Gamma$ .
- (2) For all  $x \in \text{Dom}(\sigma)$ ,  $\Box^x$  and  $\Diamond^x$  do not appear in  $\Gamma$ , thus all the occurrences of  $x$  in  $\Gamma$  are free.
- (3) All the free variables  $x$  in  $\Gamma$  are in  $\text{Dom}(\sigma)$ .
- (4) If  $\Box^x \varphi, \Box^y \psi \in \Gamma$  then  $y$  is not a free variable in  $\varphi$ .
- (5) If  $\Box^x \varphi, \Diamond^y \psi \in \Gamma$  then  $y$  is not a free variable in  $\varphi$  and  $x$  is not a free variable in  $\psi$ .
- (6) For any  $x$  in  $\text{Dom}(\sigma)$ ,  $\sigma(x) = x$ .

(1): We prove it by induction on the structure of  $\mathcal{T}$  from the root: it is true for the clean formula  $\theta$  by definition, and all the rules preserve this property since they never add any new occurrences of modalities.

(2): Again, we prove it by induction from the root: at the root the claim is true by the definition of  $\sigma_r$  and the cleanness of  $\theta$  (cf. also Footnote 12). Moreover, all the rules preserve this property. In particular, for the rule (BR), by induction hypothesis, for any variable  $x \in \text{Dom}(\sigma)$ , it only occurs free in the formulas of the numerator. Thus the occurrences of  $x \in \text{Dom}(\sigma)$ , if any, are also free in any  $\varphi_j$  and  $\psi_i[y/y_i]$  which have less modalities to bind than the numerator. Now for  $x \in \text{Dom}(\sigma') \setminus \text{Dom}(\sigma)$ , i.e.,  $\Box^x$  appears in the numerator, the statement also holds by Claim (1), since there is only one occurrence of any modality for such an  $x$  in the numerator.

(3): Again, we can show by induction:  $\text{Dom}(\sigma_r)$  has all the free variables in  $\theta$ , and all the rules preserve this property. For the rule (BR), we need to check the free variables in those  $\varphi_j$  and  $\psi_i$ . Note that the only possible extra free variable in  $\varphi_j$  but not in  $\Box^{x_j} \varphi_j$  is  $x_j$  but it is already included in  $\text{Dom}(\sigma')$ . The only possible extra free variable in  $\psi_i[y/y_i]$  but not in  $\Diamond^{y_i} \psi_i$  is  $y$  which is also already in  $\text{Dom}(\sigma')$ .

(4): Towards contradiction, suppose  $\Box^x \varphi, \Box^y \psi \in \Gamma$ , but  $y$  is a free variable in  $\varphi$ . According to Claim (3)  $y \in \text{Dom}(\sigma)$ , then by Claim (2),  $\Box^y$  does not appear in  $\Gamma$ , which is in contradiction with  $\Box^y \psi \in \Gamma$ .

(5): Similarly to (4).

(6): Obvious, by definition.

Now we are ready to prove the main theorem.

**From top to bottom:** Given an open tableau  $\mathcal{T}$  from the root node  $(r : \varphi, \sigma_r)$ , we define  $\mathcal{M} = \langle W, D, \delta, R, \rho \rangle$  where:

- $W = \{w \mid (w, \Gamma, \sigma) \text{ appears in } \mathcal{T} \text{ for some } \Gamma \text{ and } \sigma\}$
- $wRv$  iff  $v = ww'$  for some  $v'$ .
- $\delta(w) = \text{Dom}(t_w)$
- $D = \bigcup_{w \in W} \delta(w)$
- $\bar{x} \in \rho(w, P)$  iff the atomic formula  $P\bar{x}$  appears in  $t_w$ .

where  $t_w$  is the last node of  $w$  in  $\mathcal{T}$ , as defined before.

According to the rule (BR) and the definition of  $\delta$ ,  $\mathcal{M}$  is indeed an increasing domain model.<sup>13</sup> Since  $(z, z) \in \text{Dom}(\sigma_r)$ ,  $D_w$  is not empty for any  $w \in W$ . Moreover  $\rho$  is well-defined due to the openness of  $\mathcal{T}$ . Note that due to Claim (3), if the atomic formula  $P\bar{x}$  appears in  $t_w$  then  $\bar{x} \subseteq D_w$ . We will show that  $\mathcal{M}, r$  is indeed a model of  $\theta$  w.r.t.  $\sigma_r$ .

<sup>13</sup>Note that it does not mean  $D_w$  is everywhere the same due to branching nodes in  $\mathcal{T}$ .

From Claim (3), if  $(w : \Gamma, \sigma)$  appears in  $\mathcal{T}$ , then all the free variables are in  $Dom(\sigma) \subseteq D_w = Dom(t_w)$ , thus it makes sense to ask whether  $\mathcal{M}, w, \sigma \models \Gamma$  (the assignment to the bound variables are irrelevant for the truth of formulas in  $\Gamma$ ). To prove that  $\mathcal{M}, w, \sigma \models \Gamma$  for all nodes  $(w : \Gamma, \sigma)$  in  $\mathcal{T}$ , we do induction on the nodes of  $\mathcal{T}$  in a bottom-up fashion from leaf nodes, by following the the rules conversely.

Leaf For leaf nodes  $(w : l_1, \dots, l_k, \sigma)$ , by definition of  $\rho$ , the statement holds based on the fact that  $\sigma(x) = x$  for all the free variable  $x$  in those literals by Claims (3), (6).

END Supposing  $\mathcal{M}, w, \sigma \models l_1 \wedge \dots \wedge l_k$ , then it is clear that  $\mathcal{M}, w, \sigma \models \Box^{x_1} \varphi_1 \wedge \dots \wedge \Box^{x_n} \varphi_n \wedge l_1 \wedge \dots \wedge l_k$  since there is no outgoing transition from  $w$ , and  $D_w$  is not empty.

$\vee, \wedge$  Obvious.

BR Suppose  $\mathcal{M}, wv_{y_i}^y, \sigma' \models \psi_i[y/y_i] \wedge \bigwedge_1^n \varphi_j$  for every  $y \in Dom(\sigma')$  and  $i \in [1, m]$ , and  $(w : \Gamma, \sigma)$  is the branching predecessor where

$$\Gamma = \{\Box^{x_j} \varphi_j \mid j \in [1, n]\} \cup \{\Diamond^{y_i} \psi_i \mid i \in [1, m]\} \cup \{l_h \mid h \in [1, k]\}.$$

Note that  $D_w = Dom(t_w) = Dom(\sigma')$ . We need to show that  $\mathcal{M}, w, \sigma \models \Gamma$ . The  $l_h$  part is as in the case of the leaf nodes. For  $\Box^{x_j} \varphi_j$ , let  $x_j \in Dom(\sigma') = D_w$  be the witness, we need to show that at all the successors  $wv_{y_i}^y$  of  $w$ ,  $\mathcal{M}, wv_{y_i}^y, \sigma[x_j \mapsto x_j] \models \varphi_j$ . Now by Claim (4), all  $x_k$  such that  $k \neq j$  are not free in  $\varphi_j$ . From the induction hypothesis (IH) that  $\mathcal{M}, wv_{y_i}^y, \sigma' \models \varphi_j$  for each  $wv_{y_i}^y$ , and the fact that  $\sigma[x_j \mapsto x_j]$  and  $\sigma'$  agree on the free variables in  $\varphi_j$ , we know  $\mathcal{M}, wv_{y_i}^y, \sigma[x_j \mapsto x_j] \models \varphi_j$  for each  $wv_{y_i}^y$ . Thus  $\mathcal{M}, w, \sigma \models \Box^{x_j} \varphi_j$ .

As for each  $\Diamond^{y_i} \psi_i$ , we show that for each  $y \in D_w = Dom(\sigma')$ ,  $\mathcal{M}, wv_{y_i}^y, \sigma[y_i \mapsto y] \models \psi_i$ . Note that  $y$  might be not in  $Dom(\sigma)$ . By the IH,  $\mathcal{M}, wv_{y_i}^y, \sigma' \models \psi_i[y/y_i]$ . Since  $\sigma'$  assigns  $y$  to  $y$ , we just need to show that  $\sigma[y_i \mapsto y]$  and  $\sigma'$  agree on all the free variables in  $\psi_i[y/y_i]$  except  $y$ . Note that  $y_i$  is not a free variable in  $\psi_i[y/y_i]$ , therefore the only possible differences between  $\sigma[y_i \mapsto y]$  and  $\sigma'$  are about those  $x_j$  where  $x_j \neq y$ . By Claim (5), we know that  $x_j$  is not a free variable in  $\psi_i[y/y_i]$  if  $x_j \neq y$ . Therefore  $\mathcal{M}, wv_{y_i}^y, \sigma[y_i \mapsto y] \models \psi_i$  for each  $y \in D_w$ , thus  $\mathcal{M}, w, \sigma \models \Diamond^{y_i} \psi_i$  for each  $i$ .

It follows that  $\mathcal{M}, w, \sigma \models \theta$ .

Now **from bottom to top**: We just need to show that the rule applications preserve the satisfiability of the formula set. Note that for  $(\vee)$  it suffices to show one outcome node is still satisfiable. In this way, there is an open tableau since if the formula sets at the leaf nodes and branching nodes are satisfiable, then there is no contradiction among the literals. It is obvious that  $(\wedge)$  and  $(\text{END})$  preserves satisfiability, and one of the denominator of  $\vee$  preserves it too. We now show that BR also does so. Supposing  $\Gamma = \{\Box^{x_1} \varphi_1, \dots, \Box^{x_n} \varphi_n, \Diamond^{y_1} \psi_1, \dots, \Diamond^{y_m} \psi_m, \bar{l}\}$ , in some branching node  $(w : \Gamma, \sigma)$ , is satisfiable, then there is a model  $\mathcal{M}, w$  and an assignment  $\eta$  such that  $\eta(x) \in D_w$  for all  $x \in \mathbf{X}$  and:

$$\mathcal{M}, w, \eta \models \{\Box^{x_1} \varphi_1, \dots, \Box^{x_n} \varphi_n, \Diamond^{y_1} \psi_1, \dots, \Diamond^{y_m} \psi_m\} \quad (\circ)$$

By the semantics, we know there are  $a_1, \dots, a_n \in D_w$ , such that for all the successors  $v$  of  $w$ :  $\mathcal{M}, v, \eta[x_j \mapsto a_j] \models \varphi_j$ . Due to Claim (4), each  $x_j$  is not free in  $\varphi_k$  for  $k \neq j$  thus we can safely obtain for all successor  $v$  of  $w$ :

$$\mathcal{M}, v, \eta[\bar{x} \mapsto \bar{a}] \models \{\varphi_1, \dots, \varphi_n\}. \quad (A)$$

Now also by  $(\circ)$  and the semantics for  $\Diamond^{y_i}$ , for each  $\psi_i$  and each  $b \in D_w$ , there is a successor  $v_i^b$  of  $w$  such that

$$\mathcal{M}, v_i^b, \eta[y_i \mapsto b] \models \psi_i \quad (B)$$

From Claim (5),  $y_i$  is not one of  $x_j$ . Now from (A) and (B), we have for each  $\psi_i$  and each  $b \in D_w$ , there is a successor  $v_i^b$  of  $w$  such that:

$$\mathcal{M}, v_i^b, \eta[\bar{x} \mapsto \bar{a}, y_i \mapsto b] \models \{\varphi_1, \dots, \varphi_n, \psi_i\} \quad (\star)$$

Now consider any  $y \in \text{Dom}(\sigma')$  as in the denominator of the rule BR, we need to show  $\{\varphi_1, \dots, \varphi_n, \psi_i[y/y_i]\}$  is satisfiable. Suppose that  $y \in \text{Dom}(\sigma')$  and  $\eta(y) = b \in D_w$ . By Claim (2),  $y$  is free in  $\psi_i[y/y_i]$ , thus by  $(\star)$ :

$$\mathcal{M}, v_i^b, \eta[\bar{x} \mapsto \bar{a}] \models \{\varphi_1, \dots, \varphi_n, \psi_i[y/y_i]\}$$

Therefore,  $\{\varphi_1, \dots, \varphi_n, \psi_i[y/y_i]\}$  is satisfiable for each  $y \in \text{Dom}(\sigma')$ . This completes the whole proof. ■

A strong finite tree property then follows: <sup>14</sup>

**Corollary 4.1** *If an MLMS-formula  $\varphi$  (with length  $n$ ) is satisfiable, then it is satisfiable in a finite tree model  $\mathcal{M}, w$ , w.r.t. an assignment  $\sigma$  such that  $|W^{\mathcal{M}}| \leq n^{4n}$ ,  $|D^{\mathcal{M}}| \leq n$ ,  $\rho(P, w) \subseteq D_w^k$  (if  $P$  is  $k$ -ary), the depth of the tree is bounded by  $2|\varphi|$ , and  $\sigma(x) \in D_w^{\mathcal{M}}$  for each free variable  $x$  in  $\varphi$ .*

The upper bound for the depth of the tree comes from the bound on the length of the PNF of  $\varphi$ . The (very loose) upper bound on the size of  $W^{\mathcal{M}}$  comes from the fact that each node in the tableau may contain up to  $n$  modalities and each modality may have  $n$ -successors (due to the size of the domain), and the depth of the tableau is up to  $2n$ . As in normal modal logic [5, Ch. 6], we can force a binary branching tree by an MLMS-formula to get an exponential-sized model.

**Theorem 4.2** *Satisfiability problem of MLMS-formulas is PSPACE-complete.*

**Proof:**(Sketch) Note that standard modal logic formulas can be translated into our MLMS by using 0-ary predicate  $P$  for each propositional letter. For example,  $\Box\Diamond p$  can be translated as  $\Box^x\Diamond^y P$ . The PSPACE-hardness then follows since satisfiability for basic normal modal logic is PSPACE-complete [5]. For the upper bound, note that to rewrite a formula into PNF and to reletter the formula into a clean shape are efficient in space, and the length of the resulting formula is still linear in the length of the original formula. From then on, a standard PSPACE-algorithm traversing a tree structure including all the possible  $\vee$ -branches suffices, as in standard modal logic, with some care about efficiently encoding the descriptions of each node and the result of the consistent checking. ■

**Remark 1** *Our notion of the tableaux is closely related to two-player satisfiability games and alternating-tree automata [14], which will give us the algorithmic tools for MLMS in the future.*

We conjecture that a similar tableau method would work for  $\text{MLMS}^{\approx}$ , with more careful assignment management by selecting representatives of provably equivalent variables w.r.t.  $\approx$  as the local domain. We leave it to the extended version of the paper.

## 5 A new epistemic logic

In this session, we go back to the motivating epistemic setting, and give a complete axiomatization of our logic over epistemic models.

<sup>14</sup>Interested readers may go back to Footnote 6.

## 5.1 Epistemic language

To ease the presentation of the proof system, we also include the standard modality  $\Box$  as a primitive modality in the following language  $\mathbf{MLMSK}^\approx$ :<sup>15</sup>

$$\varphi ::= x \approx y \mid P\bar{x} \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid \Box^x\varphi$$

Recall that  $\mathbf{MLMSK}^\approx$  is equally expressive as  $\mathbf{MLMS}^\approx$ . In the epistemic setting, the intended reading of  $\Box\varphi$  is *i knows that  $\varphi$* , and the intended reading of  $\Box^x\varphi$  is that *i knows something such that  $\varphi$* . The semantics is as before, but we have two extra conditions on the model:

- Each  $R$  is an *equivalence relation* as in (idealized) standard epistemic logic S5. We also write  $\sim$  for  $R$ .
- For all  $w \in W$ ,  $D_w = D$ .

Note that for an increasing domain model  $\mathcal{M}$ , if  $R$  is an equivalence relation, then for any two worlds  $w, v$  such that  $wRv$ , we have  $D_w = D_v$ . Therefore, we can simply assume that there is a *constant domain* over the whole model.<sup>16</sup> We also call such models *S5-models*, following the terminology in epistemic logic.  $\varphi \in \mathbf{MLMSK}^\approx$  is *valid* if for any pointed S5-model  $\mathcal{M}, w$ , any assignment  $\sigma$ ,  $\mathcal{M}, w, \sigma \models \varphi$ . Note that in constant domain models, we no longer need to give the restriction on the assignment:  $\sigma(x) \in D = D_w$  for all assignment  $\sigma$ , all  $w \in W_{\mathcal{M}}$  and all  $x \in \mathbf{X}$ .

$\mathbf{MLMSK}^\approx$  over S5-models is a quite powerful language. As we mentioned, the semantics for  $\Box^x$  is in line with the *mention-some* reading. We can also introduce a modal operator  $\blacksquare^x$  based on *mention-all* semantics as below:

$$\mathcal{M}, w, \sigma \models \blacksquare^x\varphi \Leftrightarrow \text{for each } d \in D, \text{ either } \mathcal{M}, w, \sigma[x \mapsto d] \models \Box\varphi \text{ or } \mathcal{M}, w, \sigma[x \mapsto d] \models \Box\neg\varphi$$

Intuitively,  $\blacksquare^x\varphi$  means for all objects in the constant domain  $D$ , the agent *knows whether  $\varphi$* , e.g., *I know who came to the party* means that for each person in concern, I know whether (s)he came or not. In terms of  $\mathbf{FOML}^\approx$ ,  $\blacksquare^x\varphi$  is essentially  $\forall x(\Box\varphi \vee \Box\neg\varphi)$ . Note that  $\Box^x$  and  $\blacksquare^x$  not only differ only in the quantifiers. On the other hand, the natural  $\forall$ -version of  $\Box^x$  is defined:

$$\mathcal{M}, w, \sigma \models \Box^{\forall x}\varphi \Leftrightarrow \text{for each } d \in D, \mathcal{M}, w, \sigma[x \mapsto d] \models \Box\varphi$$

$\mathcal{M}, w, \sigma \models \Box^{\forall x}\varphi$  only says that of each object, the agent knows that it satisfies property  $\varphi(x)$ , which differs from the semantics of the *mention-all* operator.

There is also a very natural generalization of our mention-some operator with multiple variables:  $\Box^{\bar{x}}$ :

$$\mathcal{M}, w, \sigma \models \Box^{\bar{x}_1 \dots \bar{x}_n}\varphi \Leftrightarrow \text{there exist } d_1, \dots, d_n \in D \text{ such that } \mathcal{M}, t, \sigma[\bar{x} \mapsto \bar{d}] \models \Box\varphi$$

**Proposition 5.1**  $\blacksquare^x, \Box^{\forall x}$ , and  $\Box^{\bar{x}}$  can all be defined in  $\mathbf{MLMSK}^\approx$  over S5-models:

**Proof:** The following equivalences are valid for any  $\mathbf{MLMSK}^\approx$ -formula  $\varphi$ , based on the fact that  $\Diamond\Box\varphi \leftrightarrow \Box\varphi$  and  $\Box\Box\varphi \leftrightarrow \Box\varphi$  are valid on S5 models:

- $\blacksquare^x\varphi \leftrightarrow (\Diamond^x(\Box\varphi \vee \Box\neg\varphi))$
- $\Box^{\forall x}\varphi \leftrightarrow \Diamond^x\Box\varphi$
- $\Box^{\bar{x}}\varphi \leftrightarrow \Box^{x_1} \dots \Box^{x_n}\varphi$

■

<sup>15</sup>In the axiomatization we will make use of  $\Box$ . If we do not introduce them explicitly then every time we need to use  $\Box^x\varphi$  where  $x$  is not free in  $\varphi$ .

<sup>16</sup>Without the condition of constant domain, different partitions w.r.t.  $\sim$  may still have different local domains.

## 5.2 Proof system

In the rest of this section we propose a Hilbert-style proof system for the equality-free **MLMSK** and then extend it to a proof system for **MLMSK $\approx$** .

First note that the following “K axiom” for  $\Box^x$  is *not* valid on S5-models, since the witnesses object for  $x$  in  $\Box^x(\varphi \rightarrow \psi)$  and  $\Box^x\varphi$  may be different:

$$\Box^x(\varphi \rightarrow \psi) \rightarrow (\Box^x\varphi \rightarrow \Box^x\psi)$$

Therefore the modal logic of **MLMSK** is *not normal*, as expected. We propose the following proof system:

### System SMLMSK

#### Axioms

TAUT	all axioms of propositional logic
DISTK	$\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$
T	$\Box\varphi \rightarrow \varphi$
4MS	$\Box^x\varphi \rightarrow \Box\Box^x\varphi$
5MS	$\neg\Box^x\varphi \rightarrow \Box\neg\Box^x\varphi$
KtoMS	$\Box(\varphi[y/x]) \rightarrow \Box^x\varphi$ (if $\varphi[y/x]$ is admissible)
MStoK	$\Box^x\varphi \rightarrow \Box\varphi$ (if $x \notin FV(\varphi)$ )
MStoMSK	$\Box^x\varphi \rightarrow \Box^x\Box\varphi$
KT	$\Box\top$

#### Rules:

MP	$\frac{\varphi, \varphi \rightarrow \psi}{\psi}$	MONOMS	$\frac{\vdash \varphi \rightarrow \psi}{\vdash \Box^x\varphi \rightarrow \Box^x\psi}$
----	--	--------	---

where  $\varphi[y/x]$  is *admissible* if no free occurrence of  $x$  in  $\varphi$  is in the scope of a modality binding  $y$ .

The axioms are quite intuitive as they are, but they can be understood even more easily if we put them in some context to give concrete intuitive readings. For example, supposing we use  $\Box^x\varphi(x)$  to express that the agent *knows how to achieve*  $\varphi$ , then 4MS, 5MS express the (idealised) introspections of such goal-directed know-how: if you know how then you know you know how, and if you do not know how then you know you do not; KtoMS says that if you know that a particular way can achieve  $\varphi$  for sure, then you know how to achieve  $\varphi$ ; MStoK trivializes know-how to know-that, if  $x$  is irrelevant to  $\varphi$ ; MStoMSK says that if you know how to achieve  $\varphi$  then you know how to ensure that you know  $\varphi$ ; KT says we know tautologies;<sup>17</sup> and MONOMS is the monotonicity rule for know-how.

**Remark 2** *Actually, all the above axioms have incarnations in the logic system SIKH of know-how proposed in [9]. Modulo the replacement of the know-how operator Kh by  $\Box^x$ , all the axioms and rules in SIKH are derivable in SMLMSK, except the composition axiom KhKh $\rightarrow$ Kh (KhKh $\varphi \rightarrow$ Kh $\varphi$ ), which indeed captures the characteristic feature about know-how: the witness strategies can be composed.*

It is routine to verify the soundness, assuming the familiarity of first-order modal logic over S5-models.

**Theorem 5.1** *SMLMSK is sound over S5 models.*

SMLMSK is very powerful, as demonstrated by the following proposition: many usual “suspects” can be derived.

<sup>17</sup>This is a technical axiom to recover the necessitation rule for  $K$  from MONOMS. Alternatively, we can also just include the necessitation rule for  $\Box$  instead of KT.

**Proposition 5.2** *The following theorems are derivable and the rules are admissible in SMLMSK:*

MSKtoMS	$\frac{\Box^x \Box \varphi \rightarrow \Box^x \varphi}{\vdash \varphi}$	MStotK	$\Box \varphi \leftrightarrow \Box^x \varphi \ (x \notin FV(\varphi))$
NECK	$\frac{\vdash \varphi}{\vdash \Box \varphi}$	NECMS	$\frac{\vdash \varphi}{\vdash \Box^x \varphi}$
RE	$\frac{\vdash \varphi \leftrightarrow \psi}{\vdash \chi(\psi) \leftrightarrow \chi(\varphi)}$	RKtoMS	$\frac{\vdash \Box \varphi \rightarrow \psi}{\vdash \Box^x \varphi \rightarrow \psi} \ (x \notin FV(\psi))$
MST	$\Box^x \top$	RMS	$\Box^x \varphi \leftrightarrow \Box^y \varphi[y/x] \ (y \text{ is not in } \varphi)$
4	$\Box \varphi \rightarrow \Box \Box \varphi$	5	$\neg \Box \varphi \rightarrow \Box \neg \Box \varphi$

**Proof:**(Sketch) MSKtoMS is based on T and MONOMS.  $\Box \varphi \rightarrow \Box^x \varphi$  is a special case of KtoMS when  $y = x$ . Together with MStotK we have MStotK. Now from KT and MStotK we have MST. Then based on KT and MST, NECMS and NECK follow from MONOMS. The rule of Replacement of Equals (RE) can be proved inductively based on MONOMS. RKtoMS is based on MONOMS, MStoMSK, MStotK and T. The standard introspection axioms of  $\Box$  are special instances of inspection axioms 4MS and 5MS based on MStotK and RE.<sup>18</sup> Let us now look at RMS, which is the renaming axiom for bound variables. Right-to-left can be derived by starting with KtoMS:  $\vdash \Box(\varphi[y/x]) \rightarrow \Box^x \varphi$ , then since  $y$  does not appear in  $\varphi$ , we have  $\vdash \Box^y(\varphi[y/x]) \rightarrow \Box^x \varphi$  by RKtoMS. For the other direction, note that  $\varphi = (\varphi[y/x])[x/y]$  if  $y$  is not in  $\varphi$ . Then since  $(\varphi[y/x])[x/y]$  is admissible, by KtoMS,  $\vdash \Box(\varphi[y/x])[x/y] \rightarrow \Box^y \varphi[y/x]$ , namely  $\vdash \Box \varphi \rightarrow \Box^y \varphi[y/x]$ . Then by RKtoMS,  $\vdash \Box^x \varphi \rightarrow \Box^y \varphi[y/x]$ . ■

**Remark 3** *Now we know that the S5 system of  $\Box$  is a subsystem of SMLMSK. We may also view KtoMS, RKtoMS as analogues of the following axiom and rule respectively in first-order logic:*

$$\forall x \varphi \rightarrow \varphi[y/x] \quad \frac{\vdash \varphi \rightarrow \psi}{\vdash \varphi \rightarrow \forall x \psi} \ (x \notin FV(\varphi))$$

RE is the rule of replacement and RMS allows us to reletter the bound variables, as in  $\mathbf{FOL}^{\approx}$ . Moreover, the Barcan formula ( $\forall x \Box \varphi \rightarrow \Box \forall x \varphi$ ), over S5-models, can be expressed as:

$$\Diamond^x \Box \varphi \rightarrow \Box \Diamond^x \Box \varphi$$

which is also derivable in SMLMSK due to 5MS.

The following provable formula plays a role in the later completeness proof.

**Proposition 5.3**  $\vdash_{\text{SMLMSK}} (\Box^x \varphi \rightarrow \Box \psi) \rightarrow \Box(\Box^x \varphi \rightarrow \Box \psi)$ .

**Proof:** Consider the contrapositive of the formula to be proved:  $\Diamond(\Box^x \varphi \wedge \neg \Box \psi) \rightarrow (\Box^x \varphi \wedge \neg \Box \psi)$ . It is routine to show that  $\vdash \Diamond(\Box^x \varphi \wedge \neg \Box \psi) \rightarrow (\Diamond \Box^x \varphi \wedge \Diamond \neg \Box \psi)$ . Now by 5MS, 4, and T,  $\vdash \Diamond \Box^x \varphi \leftrightarrow \Box^x \varphi$  and  $\vdash \Diamond \neg \Box \psi \leftrightarrow \neg \Box \psi$ . Therefore, by RE,  $\vdash \Diamond(\Box^x \varphi \wedge \neg \Box \psi) \rightarrow (\Box^x \varphi \wedge \neg \Box \psi)$ . ■

In the following we define the Hilbert system  $\text{SMLMSK}^{\approx}$  for  $\mathbf{MLMSK}^{\approx}$  by extending SMLMSK with the following two extra axiom schemata:

#### Axioms

- ID  $x \approx x$
- SUBID  $x \approx y \rightarrow (\varphi \rightarrow \psi)$  if  $\varphi$  and  $\psi$  only differ in that some free occurrences of  $x$  in one formula are replaced by free occurrences of  $y$  in another.

<sup>18</sup>4MS can also be derived from T and 5MS in the system SMLMSK like in the case of propositional S5 system.

**Theorem 5.2**  $\text{SMLMSK}^{\approx}$  is sound.

It is routine to show:

**Proposition 5.4** The following are provable in  $\text{SMLMSK}^{\approx}$  :

$$\begin{array}{ll} \text{SYM} & x \approx y \rightarrow y \approx x \\ \text{TRANS} & x \approx y \wedge y \approx z \rightarrow x \approx z \\ \text{KEQ} & x \approx y \rightarrow \Box(x \approx y) \\ \text{KNEQ} & x \not\approx y \rightarrow \Box(x \not\approx y) \end{array}$$

**Proof:** SYM and TRANS are due to SUBID. To derive KEQ, we first have  $\vdash \Box(x \approx x)$  by ID and NECK. By SUBID we have  $\vdash x \approx y \rightarrow (\Box(x \approx y) \rightarrow \Box(x \approx x))$ . Therefore KEQ is provable.

For KNEQ, from KEQ we have  $\vdash \Box(x \not\approx y) \rightarrow x \not\approx y$ , thus  $\vdash \Box \Box x \not\approx y \rightarrow \Box x \not\approx y$  by NECK and DISTK. Note that  $\varphi \rightarrow \Box \Box \varphi$  is derivable by T and 5MS. Therefore by taking  $\varphi = x \not\approx y$  and using MP we have  $\vdash \text{KNEQ}$ . ■

## 6 Completeness

Let us first focus on the completeness of  $\text{SMLMSK}$  without axioms for equalities.

We first extend the language of  $\text{MLMSK}$  with countably infinitely many new variable symbols. Call the new language  $\text{MLMSK}^+$  and the variable set  $\mathbf{X}^+$ . In the following, we say a set of formulas is  $\text{SMLMSK}^+$ -consistent if it is  $\text{SMLMSK}$ -consistent w.r.t. the extended language  $\text{MLMSK}^+$ .

**Definition 6.1** A set of  $\text{MLMSK}^+$  formulas has  $\exists$ -property if for each  $\Box^x \varphi \in \text{MLMSK}^+$  it contains a “witness” formula  $\Box^x \varphi \rightarrow \Box \varphi[y/x]$  for some  $y \in \mathbf{X}^+$  where  $\varphi[y/x]$  is admissible.

**Definition 6.2 (Canonical model for  $\text{SMLMSK}^+$ )** The canonical model is a tuple  $\langle W^c, D^c, \sim^c, \rho^c \rangle$  where:

- $W^c$  is the set of maximal  $\text{SMLMSK}^+$ -consistent sets with  $\exists$ -property,
- $D^c = \mathbf{X}^+$ ,
- $s \sim^c t$  iff  $\Box(s) \subseteq t$  where  $\Box(s) := \{\varphi \mid \Box \varphi \in s\}$ ,
- $\bar{x} \in \rho^c(P, s)$  iff  $P\bar{x} \in s$ .

It is routine to show that  $\sim^c$  is an equivalence relation, by using axioms T, 4, and 5.

**Lemma 6.1** If  $\Box \psi \notin s \in W^c$  then there exists a  $t \in W^c$  such that  $s \sim^c t$  and  $\neg \psi \in t$ .

**Proof:** It is routine in normal modal logic to show that if  $\Box \psi \notin s$  then  $\Box(s) \cup \{\neg \psi\}$  is consistent. Now we show that  $\Box(s) \cup \{\neg \psi\}$  can be extended to an  $\text{SMLMSK}^+$ -maximal consistent set with  $\exists$ -property. We follow the general strategy in [24] by adding witness formulas one by one.

Let  $\theta_0 = \neg \psi$ . We enumerate  $\Box^x$ -formulas as:  $\Box^{x_1} \varphi_1, \Box^{x_2} \varphi_2, \dots$ . We define  $\theta_{k+1}$  as the formula:

$$\theta_k \wedge (\Box^{x_{k+1}} \varphi_{k+1} \rightarrow \Box \varphi_{k+1}[y/x_{k+1}])$$

where  $y$  is the first variable in a fixed enumeration of  $\mathbf{X}^+$  such that  $\varphi_{k+1}[y/x_{k+1}]$  is admissible and  $\Box^{x_{k+1}} \varphi_{k+1} \rightarrow \Box \varphi_{k+1}[y/x_{k+1}]$  is consistent with  $\{\theta_k\} \cup \Box(s)$ .

We now show that such  $\theta_{k+1}$  always exists (i.e., such a  $y$  exists), if  $\{\theta_k\} \cup \Box(s)$  is consistent. Towards a contradiction, suppose that  $\{\theta_k\} \cup \Box(s)$  is consistent but there is no such a  $y$ , i.e., for each  $y \in \mathbf{X}^+$  such that  $\varphi_{k+1}[y/x_{k+1}]$  is admissible there are  $\chi_1, \dots, \chi_n \in \Box(s)$  such that

$$\vdash \chi_1 \wedge \dots \wedge \chi_n \rightarrow ((\Box^{x_{k+1}} \varphi_{k+1} \rightarrow \Box \varphi_{k+1}[y/x_{k+1}]) \rightarrow \neg \theta_k).$$



By NECK and DISTK and the fact that  $\Box\chi_i \in s$  for  $1 \leq i \leq n$ , it is routine to show that  $\Box((\Box^{x_{k+1}}\varphi_{k+1} \rightarrow \Box\varphi_{k+1}[y/x_{k+1}]) \rightarrow \neg\theta_k)$  is also in  $s$ . By DISTK again,

$$\Box(\Box^{x_{k+1}}\varphi_{k+1} \rightarrow \Box\varphi_{k+1}[y/x_{k+1}]) \rightarrow \Box\neg\theta_k \quad (\star)$$

is in  $s$  for all  $y$  such that  $\varphi_{k+1}[y/x]$  is admissible.

Note that  $s$  has the  $\exists$ -property, therefore,  $\Box^{x_{k+1}}\varphi_{k+1} \rightarrow \Box\varphi_{k+1}[y^*/x]$  is in  $s$  for some particular  $y^*$  such that  $\varphi_{k+1}[y^*/x_{k+1}]$  is admissible. By Proposition 5.3,  $\Box(\Box^{x_{k+1}}\varphi_{k+1} \rightarrow \Box\varphi_{k+1}[y^*/x_{k+1}])$  is in  $s$ . By  $(\star)$ , we have  $\Box\neg\theta_k$  is in  $s$ , thus  $\neg\theta_k \in \Box(s)$  which is in contradiction with that  $\{\theta_k\} \cup \Box(s)$  is consistent.

Now since  $\{\theta_0\} \cup \Box(s) = \{\neg\psi\} \cup \Box(s)$  is consistent, we can indeed construct all the  $\theta_k$ . Let  $\Gamma = \{\theta_k \mid k \in \mathbb{N}\} \cup \Box(s)$ .  $\Gamma$  is consistent and the  $\exists$ -property is essentially built-in. We can then extend it into an SMLMSK<sup>+</sup>-maximal consistent set  $t$ . It then follows that  $s \sim^c t$  and  $\neg\varphi \in t$ . ■

Now comes the truth lemma.

**Lemma 6.2** *Let  $\sigma^*$  be the assignment such that  $\sigma^*(x) = x$  for all  $x \in X^+$ . For any  $\varphi \in \text{MLMSK}^+$ , any  $s \in W^c$ :*

$$\mathcal{M}^c, s, \sigma^* \models \varphi \Leftrightarrow \varphi \in s$$

**Proof:** We do induction on the structure of the formula  $\varphi$ .

$P\bar{x}$  By the definition of  $\rho^c$  and  $\sigma^*$ , it is obvious.

Bool Boolean cases are trivial.

$\Box\psi$  Routine, based on Lemma 6.1.

$\Box^x\psi$  Supposing  $\Box^x\psi \in s$ , by  $\exists$ -property of  $s$  there exists  $\Box\psi[y/x] \in s$  for some  $y$  such that  $\psi[y/x]$  is admissible. By 4,  $\Box\Box\psi[y/x] \in s$ . Supposing  $s \sim^c t$ , we have  $\Box\psi[y/x] \in t$  due to the definition of  $\sim^c$ . By T,  $\psi[y/x] \in t$ . By IH,  $\mathcal{M}, t, \sigma^* \models \psi[y/x]$ . Due to the definition of  $\sigma^*$  and the fact that  $\psi[y/x]$  is admissible,  $\mathcal{M}, t, \sigma^*[x \mapsto y] \models \psi$  for all  $t \sim s$ . Thus  $\mathcal{M}, s, \sigma^*[x \mapsto y] \models \Box\psi$ , it follows that  $\mathcal{M}, s, \sigma^* \models \Box^x\psi$ .

Suppose  $\Box^x\psi \notin s$ , by KtoMS,  $\Box\psi[y/x] \notin s$  for each  $y$  such that  $\psi[y/x]$  is admissible. By IH,  $\mathcal{M}^c, s, \sigma^* \models \neg\Box\psi[y/x]$  for each  $y$  such that  $\psi[y/x]$  is admissible. Due to the special assignment  $\sigma^*$  such that  $\sigma^*(y) = y$ , it is clear that  $\mathcal{M}^c, s, \sigma^*[x \mapsto y] \models \neg\Box\psi$  for each  $y$  such that  $\psi[y/x]$  is admissible.

Now consider any  $y'$  such that  $\psi[y'/x]$  is not admissible, then by RMS, we can reletter the modalities of  $y'$  in  $\psi$  with some fresh variable to obtain  $\psi'$  such that  $\vdash \psi \leftrightarrow \psi'$  and  $\psi'[y'/x]$  is admissible. Now by RE,  $\Box^x\psi' \notin s$ . We can then repeat the reasoning above to obtain  $\mathcal{M}^c, s, \sigma^*[x \mapsto y'] \models \neg\Box\psi'$  for this  $y'$ . Since  $\vdash \psi \leftrightarrow \psi'$ , by soundness,  $\mathcal{M}^c, s, \sigma^*[x \mapsto y'] \models \neg\Box\psi$ . Therefore for each  $y$ , no matter whether  $\psi[y/x]$  is admissible, we have  $\mathcal{M}^c, s, \sigma^*[x \mapsto y] \models \neg\Box\psi$ . Therefore  $\mathcal{M}^c, s, \sigma^* \models \neg\Box^x\psi$ . ■

Note that not every SMLMSK<sup>+</sup>-consistent set of formulas can be extended into a world in  $\mathcal{M}^c$ , e.g.,  $\{\Box^x\varphi(x)\} \cup \{\neg\Box\varphi(x) \mid x \in X^+\}$  cannot be extended consistently to obtain the  $\exists$ -property. However, every SMLMSK-consistent set can be extended into a world in  $W^c$  by adding the witness one by one using the new variables (cf. [24]).

**Lemma 6.3** *Every SMLMSK-consistent set of MLMSK<sup>+</sup>-formulas can be extended into an SMLMSK<sup>+</sup>-maximal consistent set of MLMSK<sup>+</sup>-formulas with  $\exists$ -property.*

Now based on Lemma 6.3 and Lemma 6.2, every SMLMSK-consistent set is satisfiable by some pointed model and an assignment. The completeness follows:

**Theorem 6.1** SMLMSK is strongly complete w.r.t. **MLMSK** over S5 models.

The completeness of  $\text{SMLMSK}^\approx$  is quite routine based on the completeness proof of SMLMSK. We only sketch the idea here following [24].

**Theorem 6.2**  $\text{SMLMSK}^\approx$  is strongly complete w.r.t.  $\text{MLMSK}^\approx$  over S5 models.

Unfortunately, **MLMSK** over (constant-domain) S5 models is indeed too powerful, we can code first-order formulas by replacing each quantifier in a first-order formula in the prenex form by  $\Box^x$  or  $\Diamond^x\Box$  respectively. We can show that this translation preserves the satisfiability. For example, we can translate  $\exists x\forall y\varphi$  ( $\varphi$  is quantifier-free) into an **MLMSK** formula  $\Box^x\Diamond^y\Box\varphi$ , which is equivalent to the first-order modal formula  $\exists x\Box\forall y\Box\varphi$  over S5 models, and it implies  $\exists x\forall y\varphi$  by reflexivity. If a first-order formula is satisfiable then we can build a single-world S5 model such that the translated **MLMSK**-formula is also satisfiable. On the other hand, if the translated **MLMSK**-formula is satisfiable in some pointed **MLMSK** S5 model, then we can just pick the designated world in that model as a first-order structure to satisfy the original first-order formula. This leads to the undecidability of **MLMSK** due to the undecidability of first-order logic:

**Theorem 6.3** **MLMSK** is undecidable over S5 models.

## 7 Future work

Due to limited space, we omit several proofs and some additional results, and leave them to the extended version of this conference paper. We believe that this is only the beginning of an interesting story. On the technical side, we may study potential properties of  $\text{MLMS}^\approx$  such as interpolation, frame definability, characterization over finite models, axiomatization and (un)decidability on various frame classes (with or without extra constant symbols). Although the full **MLMSK** over S5 models is undecidable, the concrete “propositional” know-wh logics mentioned in the introduction are usually decidable. One explanation is that the existing know-wh logics are often similar to one-variable fragments of the first-order modal language, which may lead to decidability over S5 and other models [43], e.g., the conditional *knowing value* formulas  $Kv(\varphi, c)$  discussed in [41, 42] are essentially  $\exists x\Box(\varphi \rightarrow c = x)$ . A detailed comparison with known guarded-like fragments with extra frame constraints will be very useful to understand the new framework more deeply. Moreover, we believe our techniques and results can be generalized to poly-modal (multi-agent) settings. It then makes sense to discuss the mention-some version of common knowledge operator  $C^x := \exists xC$  where  $C$  is the propositional common knowledge operator.<sup>19</sup> There is also a clear similarity with modal logic over neighbourhood models, which is worth exploring.

## References

- [1] Hajnal Andr eka, Istv an N emeti & Johan van Benthem (1998): *Modal Languages and Bounded Fragments of Predicate Logic*. *Journal of Philosophical Logic* 27(3), pp. 217–274, doi:10.1023/A:1004275029985.
- [2] Alexandru Baltag (2016): *To Know is to Know the Value of a Variable*. In: *Proceedings of AiML Vol. 11*, pp. 135–155. Available at <http://www.aiml.net/volumes/volume11/Baltag.pdf>.

<sup>19</sup>This is a rather strong notion of common knowledge, e.g., we commonly know how to prove the theorem in this sense means a fixed proof is also commonly known.

- [3] Francesco Belardinelli & Alessio Lomuscio (2011): *First-Order Linear-time Epistemic Logic with Group Knowledge: An Axiomatisation of the Monodic Fragment*. *Fundamenta Informaticae* 106(2-4), pp. 175–190, doi:10.3233/FI-2011-382.
- [4] Johan van Benthem (2010): *Frame correspondences in modal predicate logic*. In: *Proofs, categories and computations: essays in honor of Grigori Mints*, College Publications.
- [5] Patrick Blackburn, Maarten de Rijke & Yde Venema (2002): *Modal Logic*. Cambridge University Press, doi:10.1017/CBO9781107050884.
- [6] Torben Braüner & Silvio Ghilardi (2007): *First-order modal logic*. In: *Handbook of Modal Logic*, pp. 549–620, doi:10.1016/S1570-2464(07)80012-7.
- [7] Jie Fan, Yanjing Wang & Hans van Ditmarsch (2014): *Almost necessary*. In: *Proceedings of AiML Vol.10*, pp. 178–196. Available at <http://www.aiml.net/volumes/volume10/Fan-Wang-Ditmarsch.pdf>.
- [8] Jie Fan, Yanjing Wang & Hans van Ditmarsch (2015): *Contingency and Knowing Whether*. *The Review of Symbolic Logic* 8, pp. 75–107, doi:10.1017/S1755020314000343.
- [9] Raul Fervari, Andreas Herzig, Yanjun Li & Yanjing Wang (2017): *Strategically knowing how*. In: *Proceedings of IJCAI '17*. To appear.
- [10] Melvin Fitting & Richard L. Mendelsohn (1998): *First-Order Modal Logic*. Springer, doi:10.1007/978-94-011-5292-1.
- [11] Malvin Gattinger, Jan van Eijck & Yanjing Wang (2017): *Knowing Values and Public Inspection*. In: *Proceedings of ICLA'17*, pp. 77–90, doi:10.1007/978-3-662-54069-5\_7.
- [12] Paul Gochet & Pascal Gribomont (2006): *Epistemic Logic*. In Dov M. Gabbay & John Woods, editors: *Handbook of the History of Logic*, 7, doi:10.1016/S1874-5857(06)80028-2.
- [13] Rajeev Goré (1999): *Tableau Methods for Modal and Temporal Logics*. Springer Netherlands, doi:10.1007/978-94-017-1754-0\_6.
- [14] Erich Grädel, Wolfgang Thomas & Thomas Wilke, editors (2002): *Automata, Logics, and Infinite Games: A Guide to Current Research*. LNCS 2500, Springer, doi:10.1007/3-540-36387-4.
- [15] Jeroen Groenendijk & Martin Stokhof (1982): *Semantic Analysis of "Wh"-Complements*. *Linguistics and Philosophy* 5(2), pp. 175–233, doi:10.1007/bf00351052.
- [16] Tao Gu & Yanjing Wang (2016): *"Knowing value" logic as a normal modal logic*. In: *Proceedings of AiML Vol.11*, pp. 362–381. Available at [www.aiml.net/volumes/volume11/Gu-Wang.pdf](http://www.aiml.net/volumes/volume11/Gu-Wang.pdf).
- [17] Charles L Hamblin (1973): *Questions in Montague english*. *Foundations of language*, pp. 41–53, doi:10.1016/b978-0-12-545850-4.50014-5.
- [18] Helle Hvid Hansen (2003): *Monotonic Modal Logics*. Master's thesis, Universiteit van Amsterdam. Available at <https://eprints.illc.uva.nl/108/>.
- [19] Andreas Herzig & Nicolas Troquard (2006): *Knowing how to play: uniform choices in logics of agency*. In: (AAMAS 2006, pp. 209–216, doi:10.1145/1160633.1160666.
- [20] Jaakko Hintikka (1962): *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca N.Y., doi:10.1111/j.1468-0149.1963.tb00788.x.
- [21] Ian M. Hodkinson (2002): *Monodic Packed Fragment with Equality is Decidable*. *Studia Logica* 72(2), pp. 185–197, doi:10.1023/A:1021356410579.
- [22] Ian M. Hodkinson, Frank Wolter & Michael Zakharyashev (2000): *Decidable fragment of first-order temporal logics*. *Annals of Pure and Applied Logic* 106(1-3), pp. 85–134, doi:10.2307/2695115.
- [23] Ian M. Hodkinson, Frank Wolter & Michael Zakharyashev (2002): *Decidable and Undecidable Fragments of First-Order Branching Temporal Logics*. In: *Proceedings of LiCS'02*, pp. 393–402, doi:10.1109/lics.2002.1029847.
- [24] George Edward Hughes & Maxwell John Cresswell (1996): *A New Introduction to Modal Logic*. Routledge, doi:10.4324/9780203028100.

- [25] Wojciech Jamroga & Thomas Ågotnes (2007): *Constructive knowledge: what agents can achieve under imperfect information*. *Journal of Applied Non-Classical Logics* 17(4), pp. 423–475, doi:10.3166/jancl.17.423-475.
- [26] Wojciech Jamroga & Wiebe van der Hoek (2004): *Agents that Know How to Play*. *Fundam. Inform.* 63(2-3), pp. 185–219. Available at <http://content.iospress.com/articles/fundamenta-informaticae/fi63-2-3-05>.
- [27] Martin Lange (2002): *Games for Modal and Temporal Logics*. Ph.D. thesis, University of Edinburgh. Available at [www.lfcs.inf.ed.ac.uk/reports/03/ECS-LFCS-03-431/](http://www.lfcs.inf.ed.ac.uk/reports/03/ECS-LFCS-03-431/).
- [28] Tszuyuen Lau & Yanjing Wang (2016): *Knowing your ability*. *The Philosophical Forum*, pp. 415–424, doi:10.1111/phil.12130.
- [29] Yanjun Li & Yanjing Wang (2017): *Achieving while maintaining: A logic of knowing how with intermediate constraints*. In: *Proceedings of ICLA'17*, pp. 154–167, doi:10.1007/978-3-662-54069-5\_12.
- [30] Pavel Naumov & Jia Tao (2017): *Together We Know How to Achieve: An Epistemic Logic of Know-How*. In: *Proceedings of TARK'17*. Available at [arxiv.org/abs/1705.09349](http://arxiv.org/abs/1705.09349).
- [31] Marc Pauly (2000): *From Programs to Games: Invariance and Safety for Bisimulation*. In: *Proceedings of CSL'00*, Springer, pp. 485–496, doi:10.1007/3-540-44622-2\_33.
- [32] Jan Plaza (2007): *Logics of public communications*. *Synthese* 158(2), pp. 165–179, doi:10.1007/s11229-007-9168-7.
- [33] Graham Priest (2008): *An Introduction to Non-Classical Logic: From If to Is*. Cambridge University Press, doi:10.1017/cbo9780511801174.
- [34] Philipp Rohde (2002): *Expressive Power of Monadic Second-Order Logic and Modal  $\mu$ -Calculus*. In: *Automata Logics, and Infinite Games*, pp. 387–393, doi:10.1007/3-540-36387-4\_14.
- [35] Jason Stanley (2011): *Know how*. Oxford University Press, doi:10.1093/acprof:oso/9780199695362.001.0001.
- [36] Jason Stanley & Timothy Williamson (2001): *Knowing how*. *The Journal of Philosophy*, pp. 411–444, doi:10.2307/2678403.
- [37] Wiesław Szwaś & Lidia Tendera (2004): *The guarded fragment with transitive guards*. *Annals of Pure and Applied Logic* 128(1), pp. 227 – 276, doi:10.1016/j.apal.2004.01.003.
- [38] Yanjing Wang (2015): *A Logic of Knowing How*. In: *Proceedings of LORI-V*, pp. 392–405, doi:10.1007/978-3-662-48561-3\_32.
- [39] Yanjing Wang (2016): *Beyond knowing that: a new generation of epistemic logics*. In: *Jaakko Hintikka on knowledge and game theoretical semantics (forthcoming)*. Available at [arxiv.org/abs/1605.01995](http://arxiv.org/abs/1605.01995).
- [40] Yanjing Wang (2017): *A logic of goal-directed knowing how*. *Synthese*, doi:10.1007/s11229-016-1272-0. Forthcoming.
- [41] Yanjing Wang & Jie Fan (2013): *Knowing That, Knowing What, and Public Communication: Public Announcement Logic with  $K_v$  Operators*. In: *Proceedings of IJCAI'13*, pp. 1139–1146. Available at [www.aaai.org/ocs/index.php/IJCAI/IJCAI13/paper/view/6742/6782](http://www.aaai.org/ocs/index.php/IJCAI/IJCAI13/paper/view/6742/6782).
- [42] Yanjing Wang & Jie Fan (2014): *Conditionally knowing what*. In: *Proceedings of AiML Vol.10*, pp. 569–587. Available at [www.aiml.net/volumes/volume10/Wang-Fan.pdf](http://www.aiml.net/volumes/volume10/Wang-Fan.pdf).
- [43] Frank Wolter & Michael Zakharyashev (2001): *Decidable Fragments of First-Order Modal Logics*. *The Journal of Symbolic Logic* 66(3), pp. 1415–1438, doi:10.2307/2695115.
- [44] Chao Xu, Yanjing Wang & Thomas Studer (2016): *A Logic of Knowing Why*. Available at [arxiv.org/abs/1609.06405](http://arxiv.org/abs/1609.06405). Manuscript.
- [45] Reihane Zoghifard & Massoud Pourmahdian (2016): *First-Order Modal Logic: Frame Definability and Lindström Theorems*. Available at [arxiv.org/abs/1602.00201](http://arxiv.org/abs/1602.00201).