Decidability Results in First-Order Epistemic Planning

Andrés Occhipinti Liberman[1] and **Rasmus Kræmmer Rendsvig**[2]

[1]DTU Compute, Technical University of Denmark
[2]Center for Information and Bubble Studies, University of Copenhagen
aocc@dtu.dk, rasmus@hum.ku.dk

Abstract

Propositional Dynamic Epistemic Logic (DEL) provides an expressive framework for epistemic planning, but lacks desirable features that are standard in first-order planning languages (such as problem-independent action representations via action schemas). A recent epistemic planning formalism based on First-Order Dynamic Epistemic Logic (FODEL) combines the strengths of DEL (higher-order epistemics) with those of first-order languages (lifted representation), yielding benefits in terms of expressiveness and representational succinctness. This paper studies the *plan existence problem* for FODEL planning, showing that while the problem is generally undecidable, the cases of single-agent planning and multi-agent planning with non-modal preconditions are decidable.

1 Introduction

The development of autonomous agents is a key goal of artificial intelligence. A salient feature of autonomous agents is their ability to exhibit goal-directed behaviour, i.e., to commit to goals and generate plans to achieve them. In a multi-agent setting, agents often need to create their own plans taking into account, not only their own capabilities and knowledge, but also their knowledge about other agents. *Epistemic planning* [Baral *et al.*, 2017] focuses on domains where such form of reasoning is key. For example, in an epistemic planning problem, agent a may have the following goal: "a knows the truth-value of φ, while b does not know that a knows this". To achieve this goal, agent a will typically need to reason about what it can do to learn φ, but also about what b may do or learn about φ from a's actions.

In automated planning, standard formalisms such as the Planning Domain Definition Language (PDDL) model domains using first-order logic. A notable advantage of first-order over propositional languages is that variables can be used to define domain dynamics compactly. E.g., in the PDDL description of BlocksWorld, the *action schema* stack(x, y) uses variables x and y to represent the preconditions and effects of all actions of the form: "put block x on top of block y". By lifting the representation to the level of variables, schemas yield action representations whose size is independent of the number of objects in a domain.

Formalisms such as PDDL can describe how actions change the *physical* state of the environment, but rarely model the *mental* states of agents or knowledge dynamics. Automated planning has therefore recently seen an influx of work on *epistemic planning* based on Dynamic Epistemic Logic (DEL). DEL offers a highly expressive basis for planning, allowing e.g. nondeterminism, partial observability and arbitrary levels of higher-order reasoning. For an overview, see e.g. [Bolander, 2017]. However, the DEL language is *propositional*, and this has drawbacks. For instance, DEL cannot provide problem-independent action definitions via schemas.

Recently, Achen, Occhipinti Liberman and Rendsvig [2020] introduced an epistemic planning formalism based on First-Order Dynamic Epistemic Logic (FODEL). FODEL integrates the higher-order expressivity of DEL with the first-order abstraction level of languages such as PDDL. It is based on term-modal logic: first-order modal logics where the operators double as predicates by taking first-order terms as subscripts, with e.g. $\exists x K_x P(x)$ a well-formed formula. This allows agents and epistemic formulas in action schemas. While this yields advantages in terms of expressiveness and representational succinctness, the decidability and complexity aspects of FODEL have not been explored. This paper studies the decidability of the *plan existence problem* for FODEL planning.

For DEL, the problem is undecidable in general [Bolander and Andersen, 2011], entailing that the unrestricted FODEL problem is undecidable as well, since FODEL planning extends DEL planning. However, decidable and reasonably expressive fragments of DEL planning have been found, such as single-agent planning and multi-agent planning with non-modal preconditions. The main results of this paper show that *the corresponding FODEL fragments are also decidable*.

Before reaching the main results, Section 2 presents the FODEL framework: its syntax, semantics, action schemas and updates, and planning notions, while Section 3 introduces first-order bisimulations for term-modal models. The notion and accompanying results are essential to the study of FODEL planning: As for DEL planning [Bolander and Andersen, 2011; Yu *et al.*, 2013], the decidability proofs of Section 4 rely crucially on constraining the state space, to ensure that only finitely many states must be visited to check for a

plan. In the decidable DEL cases, this is achieved by showing that the infinite state space to be searched is finitely representable *up to bisimulation*. Hence, the notion of bisimulation is pivotal to such proofs. Introducing FODEL bisimulations and developing the accompanying model-theoretic results allows us to apply similar arguments to FODEL planning and ultimately establish decidability for single-agent planning as well as multi-agent planning with non-modal preconditions.

2 First-Order Dynamic Epistemic Logic

Throughout, we adopt the epistemic planning framework developed in detail in [Achen *et al.*, 2020]. Additional FODEL examples may be found in [Occhipinti Liberman and Rendsvig, 2019]. For an introduction to epistemic planning with DEL, see e.g. [Bolander, 2017].

2.1 Language

Term-modal logical languages are parameterized by a *signature* specifying the *non-logical symbols* and their *type*—i.e., the constants and relation symbols, and the sort of arguments (agents or objects) they apply to. For the purposes of representing epistemic planning, we limit attention to function-free signatures with finitely many predicate and constant symbols. This restriction is not assumed in [Achen *et al.*, 2020].

Definition 1. A signature *is a tuple* $\Sigma = (V, C, R, t)$ *where* $V := V_{\mathtt{agt}} \cup V_{\mathtt{obj}}$ *is a set of agent and object variables, with* $V_{\mathtt{agt}}$ *and* $V_{\mathtt{obj}}$ *both countably infinite,* C *a finite set of constant symbols and* R *a finite set of predicate symbols. The* terms *of* Σ *are* $T := V \cup C$. *Finally,* t *is a type assignment that satisfies (i) for* $x \in T$, $t(x) \in \{\mathtt{agt}, \mathtt{obj}\}$ *with* $t(x) = i$ *for all* $x \in V_i$, $i \in \{\mathtt{agt}, \mathtt{obj}\}$, *and (ii) for* $r \in R$ *with arity* n, $t(r) \subseteq \{\mathtt{agt}, \mathtt{obj}\}^n$. *We denote by* $t_i(r)$ *the* i-th component *of* $t(r)$ *and by* $ar(r)$ *the arity of* r.

Definition 2. *Let* Σ *be a signature. The* language $\mathcal{L}(\Sigma)$ *is given by the grammar:*

$$\varphi := r(t_1, ..., t_n) \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_t\varphi \mid \forall x\varphi$$

where $t_1, \ldots, t_n \in T$, $r \in R$, $(t(t_1), ..., t(t_n)) \in t(r)$, $t \in T$ *with* $t(t) = \mathtt{agt}$, *and* $x \in V$. *An* atom *is a formula of the form* $r(t_1, ..., t_n)$; *it is* ground *if it contains no variables and* free *if it only contains variables.* $\mathtt{GroundAtoms}(\mathcal{L}(\Sigma))$ *and* $\mathtt{FreeAtoms}(\mathcal{L}(\Sigma))$ *denote the set of ground and free atoms of* $\mathcal{L}(\Sigma)$. *A* sentence *is a formula without free variables.* $\mathcal{L}_0(\Sigma)$ *is the* K_t-free sublanguage.

Throughout, the standard Boolean connectives as well as $\top$, $\bot$ and $\exists$ are used as meta-linguistic abbreviations as usual. $K_t\varphi$ is read as "agent t knows that φ".

2.2 Epistemic Models

In epistemic logic-based planning frameworks, states are represented using variants of *Kripke models* [Kripke, 1962]. We use *constant domain* Kripke models (the same domain in each world) with *non-rigid* constants (names, like predicates and relations, may change extension between worlds, allowing for uncertainty about agent identity). Models where agents' accessibility relations are equivalence relations are standard in epistemic planning, and here called *epistemic models*. A
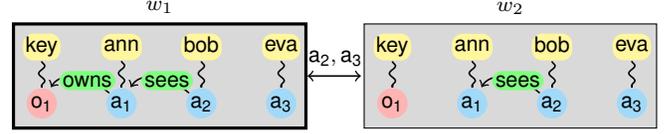

Figure 1: An epistemic state $s_0 = (M, w_1)$ for $\mathcal{L}(\{x^\star, x_1, x_2, \ldots\}, \{\mathsf{ann}, \mathsf{bob}, \mathsf{eva}, \mathsf{key}\}, \{\mathsf{owns}, \mathsf{sees}\}, t)$. Types are $t(c) = \mathtt{agt}$ for $c \in \{\mathsf{ann}, \mathsf{bob}, \mathsf{eva}\}$, $t(\mathsf{key}) = \mathtt{obj}$, $t(\mathsf{owns}) = (\mathtt{agt}, \mathtt{obj})$ and $t(\mathsf{sees}) = (\mathtt{agt}, \mathtt{agt})$. The state has worlds $W^M = \{w_1, w_2\}$, domain $D^M = \{\mathsf{a_1, a_2, a_3, o_1}\}$ and accessibility relations $R_{\mathsf{a_1}} = \{(w_1, w_1), (w_2, w_2)\}$, $R_{\mathsf{a_2}} = R_{\mathsf{a_3}} = W \times W$ (reflexive world-edges are not drawn). The interpretation function is depicted like so: squiggly lines link constants to entities, and predicate-labelled arrows relate entities for predicates (e.g., $(\mathsf{a_1}, \mathsf{o_1}) \in I^M(\mathsf{owns}, w_1)$). In the actual world w_1, ann owns the key. bob and eva do not know this (they cannot distinguish w_1 and w_2). Everyone knows that bob can see ann.

model with a designated world is a *state*; Figure 1 depicts an *epistemic state*, formally defined below.

Definition 3. *An* $\mathcal{L}(\Sigma)$ model *is a tuple* $M = (D^M, W^M, R^M, I^M)$ *where (i) the* domain $D^M := D_{\mathtt{agt}} \dot\cup D_{\mathtt{obj}}$ *is the disjoint union of the non-empty sets* $D_{\mathtt{agt}}$ *and* $D_{\mathtt{obj}}$ *(the* agent domain *and the* object domain*); (ii)* W^M *is a non-empty set of* worlds*; (iii)* $R^M : D_{\mathtt{agt}} \to \mathcal{P}(W \times W)$ *associates to each agent* $d \in D_{\mathtt{agt}}$ *an accessibility relation on* W, *denoted* R_d^M, *and (iv)* I^M *is an* interpretation *satisfying:* $I^M(c, w) \in D_{t(c)}$ *if* $c \in C$ *and* $I^M(r, w) \subseteq \prod_{i=1}^n D_{t_i(r)}$ *for* $r \in R$ *with arity* n. *For* $w \in W^M$, *the pair* $s = (M, w)$ *is called a* state. *A state* s *is called* finite *if* W^M *is finite. If* R_d^M *is an equivalence relation for all* $d \in D_{\mathtt{agt}}$, *then* M *is an* epistemic model, *and* $s = (M, w)$ *is then an* epistemic state.

Truth conditions for formulas of $\mathcal{L}(\Sigma)$ are defined over states, together with a variable assignment.

Definition 4. *A* variable assignment *for a model* M *is a map* $g : V \to D^M$ *with* $g(x) \in D_{t(x)}$ *for each* $x \in V$. *Denote by* $g[x \mapsto d]$ *the assignment mapping* x *to* d *but else identical to* g. *The* extension *of term* $t \in T$ *at* w *in* M, *under* g, *is* $\llbracket t \rrbracket_w^{I^M, g} = g(t)$ *if* $t \in V$ *and* $\llbracket t \rrbracket_w^{I^M, g} = I^M(t, w)$ *if* $t \in C$. *If* M *is given by context, we omit it and write* $\llbracket t \rrbracket_w^{I, g}$.

Definition 5. *Let* M *and* g *be given. Then*
$M, w \vDash_g r(t_1, ..., t_n)$ *iff* $(\llbracket t_1 \rrbracket_w^{I, g}, ..., \llbracket t_n \rrbracket_w^{I, g}) \in I^M(r, w)$
$M, w \vDash_g \neg\varphi$ *iff not* $M, w \vDash_g \varphi$.
$M, w \vDash_g \varphi \wedge \psi$ *iff* $M, w \vDash_g \varphi$ *and* $M, w \vDash_g \psi$.
$M, w \vDash_g \forall x\varphi$ *iff* $M, w \vDash_{g[x \mapsto d]} \varphi$ *for all* $d \in D_{t(x)}$.
$M, w \vDash_g K_t\varphi$ *iff* $M, v \vDash_g \varphi$ *for all* v *s.t.* $(w, v) \in R_{\llbracket t \rrbracket_w^{I, g}}^M$.

2.3 Actions Schemas and Actions

The FODEL planning framework uses *action schemas* to provide compact, problem-independent domain descriptions in the spirit of PDDL (see Figure 2). FODEL action schemas use variables as *action parameters*, so a single schema may be used, e.g., to represent all actions of the form 'agent x gives agent y object z', where x, y and z are agent or object variables, later instantiated into specific constants to define
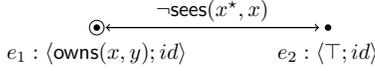
Figure 2: announce_own(x, y), the schema for agent x announcing that it owns y. Each event is drawn with the format $\langle \text{pre}(e), \text{post}(e) \rangle$, where id is the identity map (i.e., no factual change). The actual event e_1 (marked with a double circle) describes the announcement that x owns y. e_2 describes the situation in which nothing happens. Agents that do not see x will not learn what is being announced. The edge-condition $Q(e_1, e_2) = Q(e_2, e_1) = \neg\text{sees}(x^\star, x)$ defines this ($Q(e_1, e_1) = Q(e_2, e_2) = \top$ not drawn).

Figure 3: The epistemic state $s_0 \otimes$ announce_own(ann, key). bob now knows that ann owns the key. However, eva still does not know this, as she did not see bob making the announcement.

a (ground, epistemic) *action*. As in classical planning formalisms such as PDDL, a major reason for using schemas is that they allow action representations whose size is independent of the number of agents and objects in a domain.

FODEL action schemas include *preconditions* [Baltag *et al.*, 1998], *postconditions* [van Benthem *et al.*, 2006; van Ditmarsch *et al.*, 2005] as well as *edge-conditions* in the style of [Bolander, 2014]. Preconditions specify when an event is executable. Postconditions describe the physical effects of events. Edge-conditions describe how an agent observes the action in terms of the agent's circumstances.

Definition 6. *An* action schema *is a form* $a(\vec{x}) = (E, Q, \text{pre}, \text{post})$ *where (i)* a *is the* action name *and* $\vec{x} \in V^n$ *is a finite* parameter list*; (ii)* E *is a non-empty, finite set of* events*; (iii)* $Q : (E \times E) \to \mathcal{L}(\Sigma)$ *is an* edge-condition function*, where the formula* $Q(e, e')$ *has a free variable* $x^\star$ *of type* agt*, and possibly other free variables all in* $\vec{x}$*; (iv)* pre : $E \to \mathcal{L}(\Sigma)$ *assigns to each event a* precondition *formula with all free variables in* $\vec{x}$*; (v)* post : $E \to (\text{FreeAtoms}(\mathcal{L}) \rightharpoonup \mathcal{L}(\Sigma))$ *assigns to each event a* partial postcondition function *such that if* $y_1, \ldots, y_m \in \vec{x}$*, then* $\text{post}(e)(r(y_1, \ldots, y_m))$ *has all free variables from* $\vec{x}$*; else,* $\text{post}(e)(r(y_1, \ldots, y_m))$ *is undefined.* $\text{dom}(\text{post}(e))$ *denotes the set of atoms for which* $\text{post}(e)(r(t_1, \ldots, t_k)) \neq r(t_1, \ldots, t_k)$ *and is required finite for each* $e \in E$*, so postconditions are ensured finite objects.*

Schema instantiation is defined as follows.

Definition 7. *Let* $a(x_1, \ldots, x_n)$ *be an action schema and* $\sigma : \{x_1, \ldots, x_n\} \to C$ *be a* grounding substitution *mapping action parameters to constants. For* $\varphi \in \mathcal{L}(\Sigma)$*, let* $\varphi\sigma$ *be the result of replacing each free occurence of* $y \in V$ *in* φ *by* $\sigma(y)$*.*

For an action schema $a(\vec{x}) = (E, Q, \text{pre}, \text{post})$*, the* action model *induced by the grounding substitution* σ *is the tuple* $A = a(\sigma(\vec{x})) = (E^A, Q^A, \text{pre}^A, \text{post}^A)$ *where (i)* $E^A = E$ *(ii)* $Q^A(e, e') = Q(e, e')\sigma$*; (iii)* $\text{pre}^A(e) = \text{pre}(e)\sigma$*; (iv)* $\text{post}^A(e)(r(t_1, \ldots, t_n)\sigma) = \text{post}(e)(r(t_1, \ldots, t_n))\sigma$ *if* $\text{post}(e)(r(t_1, \ldots, t_n))$ *is defined; and* $\text{post}^A(e)(r(t_1, \ldots, t_n)\sigma) = r(t_1, \ldots, t_n)\sigma$ *otherwise. A pair* $a = (A, e)$ *with* $e \in E^A$ *is called an* epistemic action*, where* e *is the actual event taking place on execution.*

2.4 Product Update

Execution is defined by the operation $\otimes$ that computes the state $s' = s \otimes a$ reached by applying action a in state s. The operation is a first-order variant of *product update* from [Baltag *et al.*, 1998]. The new state s' represents the agents perception of $s = (M, w)$ after the occurrence of the action
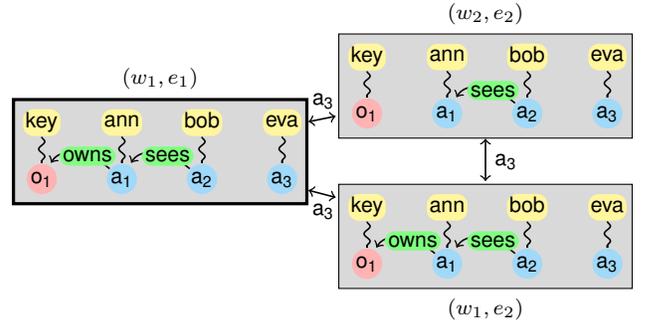
described by $a = (A, e)$. Figure 3 exemplifies the product update of the state from Figure 1 and an instance of the schema in Figure 2.

Definition 8. *Let* $a = ((E, Q, \text{pre}, \text{post}), e)$ *be an action and* $s = ((D, W, R, I), w)$ *a state. The* product update *of* s *and* a *is the state* $s \otimes a = ((D, W', R', I'), (w, e))$ *where for any* g

- $W' = \{(w, e) \in W \times E \mid (M, w) \vDash_g \text{pre}(e)\}$

- $(w, e)R'_d(w', e')$ *iff* wR_dw' *and* $M, w \vDash_{g[x^\star \mapsto d]} Q(e, e')$

- $I'(c, (w, e)) = I(c, w)$ *for all* $c \in C$ *and*
  $I'(r, (w, e)) = (I(r, w) \cup r^+(w)) \setminus r^-(w)$*, where*
  $$r^+(w) := \{(\llbracket t_1 \rrbracket_w^{I,g}, \ldots, \llbracket t_k \rrbracket_w^{I,g}) \mid$$
  $$(M, w) \vDash_g \text{post}(e)(r(t_1, \ldots, t_k))\}$$
  $$r^-(w) := \{(\llbracket t_1 \rrbracket_w^{I,g}, \ldots, \llbracket t_k \rrbracket_w^{I,g}) \mid$$
  $$(M, w) \nvDash_g \text{post}(e)(r(t_1, \ldots, t_k))\}$$

An action $a = (A, e)$ *is* applicable *in a state* $s = (M, w)$ *iff* $(M, w) \vDash_g \text{pre}(e)$*. Else* $s \otimes a$ *is undefined.*

2.5 Planning Notions

As presented in [Ghallab *et al.*, 2004], any *classical planning domain* can be described as a state-transition system $T = (\mathcal{S}, \mathcal{A}, \gamma)$ where $\mathcal{S}$ is a finite or recursively enumerable set of finite states, $\mathcal{A}$ is a finite set of actions and $\gamma : \mathcal{S} \times \mathcal{A} \rightharpoonup \mathcal{S}$ is a partial, computable state-transition function. A *classical planning task* is given by a triple (T, s_0, S_G), where T is a state-transition system, $s_0 \in \mathcal{S}$ is the initial state and $S_G \subseteq \mathcal{S}$ is the set of goal states. A *solution* to a classical planning task (T, s_0, S_G) is a plan consisting of a finite sequence of actions $a_1, a_2, \ldots, a_n$ such that (1) For all $i \leq n$, $\gamma(\gamma(\ldots \gamma(\gamma(s_0, a_1), a_2), \ldots, a_{i-1}), a_i)$ is defined, and (2) $\gamma(\gamma(\ldots \gamma(\gamma(s_0, a_1), a_2), \ldots, a_{n-1}), a_n) \in S_G$.

Epistemic planning tasks can be defined as special cases of classical planning tasks.

Definition 9. *Let* A *be a finite set of* $\mathcal{L}(\Sigma)$ *action schemas. An* epistemic planning task *based on* A *is a triple* $P = (s_0, \mathcal{A}, \varphi_G)$ *where the initial state* s_0 *is a finite epistemic state with a finite domain,* $\mathcal{A}$ *is the set of all ground instances of the schemas in* A*, and the goal formula* φ_G *is a sentence of* $\mathcal{L}(\Sigma)$*. Any epistemic planning task* $(s_0, \mathcal{A}, \varphi_G)$ *induces a classical planning task* $((\mathcal{S}, \mathcal{A}, \gamma), s_0, S_G)$ *given by:*

- $\mathcal{S} := \{s_0 \otimes a_1 \otimes \cdots \otimes a_n \mid n \in \mathbb{N}, a_i \in \mathcal{A}\}$

- $S_G := \{s \in \mathcal{S} \mid s \vDash \varphi_G\}$

- $\gamma(s,a) := s \otimes a$ if a is applicable in s, else undefined.

A solution *to an epistemic planning task is a solution to the induced classical planning task.*

We follow Aucher and Bolander [2013] in defining the plan existence problem:

Definition 10. *Let $n \in \mathbb{N}$.* PlanEx(n) *is the problem: "Given an epistemic planning task $P = (s_0, \mathcal{A}, \varphi_G)$ where s_0 is an n-agent $\mathcal{L}(\Sigma)$-epistemic state, does P have a solution?".*

3 First-Order Bisimulations

To obtain a bisimulation notion for first-order, term-modal models, we retain the "back" and "forth" conditions of propositional modal logical bisimulation, but replace the "atomic harmony" condition with first-order logical isomorphism. The resulting notion implies modal equivalence and is preserved under product updates. Further, n-*bisimulation* is defined and shown to imply modal equivalence for formulas of restricted modal depth, and states are shown to be bisimilar to their *bisimulation contractions*. These results all extend to FODEL standard properties of bisimulations in propositional DEL and are key in proving the later decidability proofs, as they allow us to show that the state spaces for some planning fragments are finitely representable, up to bisimulation.[1]

3.1 Isomorphisms, Bisimulations and Equivalence

We first recall the notion of an isomorphism, viewed here as a relationship between worlds in $\mathcal{L}(\Sigma)$ models.

Definition 11. *Let M and N be $\mathcal{L}(\Sigma)$ models. An isomorphism between $w \in W^M$ and $v \in W^N$ is a bijection $f : D^M \to D^N$ such that, for all $c \in \mathtt{C}$ and $r \in \mathtt{R}$,*

1. $f(I^M(c,w)) = I^N(c,v)$

2. $(d_1, ..., d_i) \in I^M(r,w)$ iff $(f(d_1), ..., f(d_i)) \in I^N(r,v)$

If w and v are isomorphic *based on f, we write $w \simeq_f v$.*

We then use isomorphisms to ensure that worlds linked by a bisimulation agree on all $\mathcal{L}_0(\Sigma)$ formulas, while also relating agents—and hence accessibility relations—appropriately in the 'forth' and 'back' conditions:

Definition 12. *For models M and N, let $f : D^M \to D^N$ be a bijection. A non-empty relation $Z_f \subseteq W^M \times W^N$ is an f-based bisimulation between M and N if*

Atoms: *If $(w,w') \in Z_f$, then $w \simeq_f w'$,*

Forth: *If $(w,w') \in Z_f$ and $R_d^M wv$, then there exists v' such that $R_{f(d)}^N w'v'$ and $(v,v') \in Z_f$, for all $d \in D_{\mathtt{agt}}$,*

---

[1]The existing work on bisimulations for first-order modal logic models [van Benthem, 2010a; van Benthem, 2010b; Zoghifard and Pourmahdian, 2018] concern *varying domain* models, with their added complexity. See [Achen *et al.*, 2020] for a planning related discussion. The resulting potential isomorphism-based *world-object bisimulations* are both more general and more complex than what is needed for present purposes, while not tuned to term-modal logics.

Back: *If $(w,w') \in Z_f$ and $R_{f(d)}^N w'v'$, then there exists v such that $R_d^M wv$ and $(v,v') \in Z_f$, for all $d \in D_{\mathtt{agt}}$.*

If $(w,w') \in Z_f$ for any f-based bisimulation Z_f between M and N, call (M,w) and (N,w') bisimilar, written $(M,w) \underline{\leftrightarrow} (N,w')$.

Since any two bisimilar worlds w and w' are required to be isomorphic, the relation $(M,w) \underline{\leftrightarrow} (N,w')$ can only hold between models with domains of equal cardinality. Next, we define n-bisimulations and bisimulation contractions.

Definition 13. *For models M and N, let $f : D^M \to D^N$ be a bijection. A non-empty relation $Z_f^n \subseteq W^M \times W^N$ is an f-based n-bisimulation between states (M,w) and (N,w') if $w \simeq_f w'$, and either $n = 0$ or, for some f-based $(n-1)$-bisimulation Z_f^{n-1}*

Forth: *If $R_d^M wv$ then there exists v' such that $R_{f(d)}^N w'v'$ and $v Z_f^{n-1} v'$, for all $d \in D_{\mathtt{agt}}$,*

Back: *If $R_{f(d)}^N w'v'$ then there exists v such that $R_d^M wv$ and $v Z_f^{n-1} v'$, for all $d \in D_{\mathtt{agt}}$.*

If $(w,w') \in Z_f^n$ for any f-based n-bisimulation Z_f^n between (M,w) and (N,w'), call them n-bisimilar, written $(M,w) \underline{\leftrightarrow}_n (N,w')$.

Definition 14. *The* bisimulation contraction *of a $\mathcal{L}(\Sigma)$ model M is $M_{\underline{\leftrightarrow}} = (D^M, W', R', I')$ where*

- $W' = \{[w]_{\underline{\leftrightarrow}} \mid w \in W^M\}$ where $[w]_{\underline{\leftrightarrow}} = \{v \in W^M \mid (M,w) \underline{\leftrightarrow} (M,v)\}$;

- $([w]_{\underline{\leftrightarrow}}, [v]_{\underline{\leftrightarrow}}) \in R_d'$ iff there are $w' \in [w]_{\underline{\leftrightarrow}}$ and $v' \in [v]_{\underline{\leftrightarrow}}$ such that $R_d^M w'v'$;

- For all $w \in W$, $c \in \mathtt{C}$ and $r \in \mathtt{R}$, $I'(c, [w]_{\underline{\leftrightarrow}}) = I^M(c,w)$ and $I'(r, [w]_{\underline{\leftrightarrow}}) = I^M(r,w)$

The n-bisimulation contraction $M_{\underline{\leftrightarrow}_n}$ is defined in the same way based on $\underline{\leftrightarrow}_n$.

Finally, we fix the relevant notions of logical equivalence:

Definition 15. *Let M and N be models. Two worlds $w \in W^M$ and $v \in W^N$ are* elementary equivalent, *if for any sentence $\varphi \in \mathcal{L}_0(\Sigma)$ $M, w \vDash_g \varphi$ iff $N, v \vDash_g \varphi$. If the same holds for any sentence $\varphi \in \mathcal{L}(\Sigma)$, then they are* modally equivalent.

3.2 Results on First-Order Bisimulations

With the notions defined, we turn to showing the results needed for the later decidability proofs. First, we recall a well-known result (see, e.g. [Rothmaler, 2018]).

Lemma 1. *If w and v worlds in M and N respectively, then $w \simeq_f v$ implies that w and v are elementary equivalent. If M and N have finite domains, then the converse also holds.*

From this, we show that bisimulation implies modal equivalence, without restriction. For the converse, call a model M *image-finite* models if for each world $w \in W^M$ and each relation R_d^M, the set $\{v \in W^M \mid R_d^M wv\}$ is finite. We then establish a variant of the Hennessy-Milner Theorem known from propositional modal logic:

Proposition 1. *1. If there is an f-based bisimulation between (M, w) and (N, w') then for all $\varphi \in \mathcal{L}(\Sigma)$: $M, w \vDash_g \varphi$ iff $N, w' \vDash_{f \circ g} \varphi$. 2. If M and N are image-finite and with finite domains, the converse also holds.*

Proof. 1. By induction on the complexity of φ. The base case where φ is an atomic formula follows from Lemma 1. Boolean cases are straightforward. We show the modal case with $\varphi = K_t \psi$. If $t \in \mathtt{V}$, assume that $M, w \vDash_g K_t \psi$ and $R_{g(t)}^M wv$. As $w \underline{\leftrightarrow} w'$, there is a v' such that $R_{f(g(t))}^N w'v'$ and $v \underline{\leftrightarrow} v'$. By the induction hypothesis, we get $N, v' \vDash_{f \circ g} \varphi$. As v was picked arbitrarily, $M', w' \vDash_{f \circ g} K_t \psi$. If instead $t \in \mathtt{C}$, assume that $M, w \vDash_g K_t \psi$ and $R_{I(t,w)}^M wv$. As $w \underline{\leftrightarrow} w'$, there is a v' such that $R_{f(I^M(t,w))}^N w'v'$ and $v \underline{\leftrightarrow} v'$. By the induction hypothesis, we get $N, v' \vDash_{f \circ g} \varphi$. Given $w \underline{\leftrightarrow} w'$ we get $w \simeq_f w'$ and thus $I^N(t, w') = f(I^M(t, w))$. Hence, since v was arbitrary and $I^N(t, w') = f(I^M(t, w))$, we get $N, w' \vDash_{f \circ g} K_t \varphi$. The other direction is similar.
2. The modal equivalence relation may be shown to be a bisimulation. The proof is analogous to the propositional case (see [Blackburn *et al.*, 2001, Thm 2.24]), save for Atoms. But this holds as modal equivalence of worlds joint with the finite domain assumption by Lemma 1 implies isomorphism. $\square$

The modal depth of a formula φ, denoted $md(\varphi)$, is the depth of nesting of modal operators in φ. As in the propositional case, n-bisimilar states agree on all formulas of $\mathcal{L}(\Sigma)$ whose modal depth is at most n.

Proposition 2. *If there is an f-based n-bisimulation between (M, w) and (N, w'), then for all $\varphi \in \mathcal{L}(\Sigma)$ with $md(\varphi) \leq n$: $M, w \vDash_g \varphi$ iff $N, w' \vDash_{f \circ g} \varphi$.*

Proof. By induction on n. For $n = 0$, note that $w \simeq_f v$. Thus by Lemma 1, w and w' are elementary equivalent, establishing the correspondence for sentences of modal depth 0. When free variables are assigned using g and $f \circ g$, this extends to all 0-depth formulas. The inductive step is standard (see, e.g. [Blackburn *et al.*, 2001]), except for checking that the interpretations of modal indices match appropriately between models, which they do, given f. $\square$

As in the propositional case, any state is bisimilar to its bisimulation contraction, a result used in the second decidability proof below.

Proposition 3. *For any (M, w), $(M, w) \underline{\leftrightarrow} (M_{\underline{\leftrightarrow}}, [w]_{\underline{\leftrightarrow}})$ and for any j such that $j \geq n \geq 0$, $(M, w) \underline{\leftrightarrow}_n (M_{\underline{\leftrightarrow}_j}, [w]_{\underline{\leftrightarrow}_j})$.*

Proof. The proof for each case is analogous to its propositional counterpart, see e.g. [Blackburn and van Benthem, 2007] for the $\underline{\leftrightarrow}$ case, and [Yu *et al.*, 2013] for the $\underline{\leftrightarrow}_n$ case. Both are by induction on n, both with induction steps that carry over to FODEL. The base cases follow by definition of $\underline{\leftrightarrow}$ and $\underline{\leftrightarrow}_j$. $\square$

Additionally towards proving decidability, we show that the variant of product update preserves bisimilarity. This is known for product update on propositional models, and here it is extended to first-order, edge-conditioned updates.

Proposition 4. *For any action (A, e), if $(M, w) \underline{\leftrightarrow} (M', w')$, then 1. $(M, w) \otimes (A, e)$ is defined iff $(M', w') \otimes (A, e)$ is, and 2. if defined, then $(M, w) \otimes (A, e) \underline{\leftrightarrow} (M', w') \otimes (A, e)$.*

Proof. 1. Follows by Proposition 1. 2. $(M, w) \underline{\leftrightarrow} (M', w')$ implies there is a bijection $f : D^M \to D^{M'}$ and an f-based bisimulation Z_f with $wZ_f w'$. Define $Z'_f \subseteq W^{M \otimes A} \times W^{M' \otimes A}$ by $(v, e_1) Z'_f (v', h_1)$ iff $vZ_f v'$ and $e_1 = h_1$. Then Z'_f is an f-based bisimulation: **Atoms**: $vZ_f v'$ requires that $v \simeq_f v'$. This implies that $(v, e) \simeq_f (v', e)$ for any e, as post will have identical effects on v and v'. **Forth**: Assume that $(v, e_1) Z'_f (v', e_1)$ and $(v, e_1) R_d^{M \otimes A} (u, e_2)$. Then $vZ_f v'$, $vR_d^M u$ and $M, v \vDash_{g[x^\star \mapsto d]} Q(e_1, e_2)$. As $vZ_f v'$, also $M', v' \vDash_{g[x^\star \mapsto d]} Q(e_1, e_2)$. As $vR_d^M u$ and $vZ_f v'$, there is u' s.t. $v'R_d^{M'} u'$ and $uZ_f u'$. As $(u, e_2) \in W^{M \otimes A}$ and $uZ_f u'$, also $(u', e_2) \in W^{M' \otimes A}$. Hence also $(v', e_1) R_d^{M' \otimes A} (u', e_2)$, so $(u, e_2) Z'_f (u', e_2)$. The proof for **Back** is similar. $\square$

A corresponding result holds as well for n-bisimulations and actions with non-modal preconditions:

Proposition 5. *For any two n-bisimilar $\mathcal{L}(\Sigma)$ states (M, w) and (M', w') and any action (A, e) with $\mathsf{pre}(e') \in \mathcal{L}_0(\Sigma)$ for all $e' \in E^A$, 1. $(M, w) \otimes (A, e)$ is defined iff $(M', w') \otimes (A, e)$ is, and 2. if defined, $(M, w) \otimes (A, e) \underline{\leftrightarrow}_n (M', w') \otimes (A, e)$.*

Proof. 1. Since $(M, w) \underline{\leftrightarrow}_n (M', w')$, $w \simeq_f w'$. As $\mathsf{pre}(e) \in \mathcal{L}_0(\Sigma)$, $M, w \vDash_g \mathsf{pre}(e)$ iff $M', w' \vDash_{f \circ g} \mathsf{pre}(e)$. 2. Assume that (A, e) is applicable in (M, w). We can prove $(M, w) \underline{\leftrightarrow}_n (M', w')$ by induction on n. Base case: $n = 0$. Since $(M, w) \underline{\leftrightarrow}_0 (M', w')$, then $w \simeq_f w'$. Thus, by the semantics of product update, $(w, e) \simeq_f (w', e)$. The inductive step is analogous to [Yu *et al.*, 2013, Proposition 4.14]. $\square$

4 Decidability Results

4.1 Single-Agent Planning

We show first that single-agent FODEL planning is decidable. The gist of the proof is that each single-agent epistemic state is shown bisimilar to a 'canonical state', and that there are finitely many such canonical states. Hence the state space for single-agent planning is finitely representable, and it can therefore be decided whether such task is solvable. We define these canonical states as follows:

Definition 16. *Given a signature Σ and a domain D with $D_{\mathtt{agt}} = \{d\}$, define the structure $\mathcal{M}_{(\Sigma, D)}$ as the tuple $(D, \mathcal{W}, \mathcal{W} \times \mathcal{W}, \mathcal{I})$ where $\mathcal{W}$ is the set of first-order interpretations i of Σ over D and $\mathcal{I}(x, i) = i(x)$ for all $x \in \mathtt{C} \cup \mathtt{R}$ and all $i \in \mathcal{W}$. The $\mathcal{M}_{(\Sigma, D)}$ substate of a single-agent $\mathcal{L}(\Sigma)$ epistemic state $(M, w) = (D, W, R, I, w)$ is the epistemic state $(M^\star, w^\star) = (D, W^\star, R^\star, I^\star, w^\star)$ where*

$$W^\star = \{i \in \mathcal{W} \mid \exists u \in W \text{ s.t. } R_d wu \text{ and } \forall x \in \mathtt{C} \cup \mathtt{R},$$
$$I(x, u) = \mathcal{I}(x, i)\},$$
$$R_d^\star = W^\star \times W^\star,$$
$$I^\star(x, i) = \mathcal{I}(x, i), \forall x \in \mathtt{C} \cup \mathtt{R}, \forall i \in W^\star,$$
$$w^\star = i^\star \text{ s.t. } \mathcal{I}(x, i^\star) = I(x, w), \forall x \in \mathtt{C} \cup \mathtt{R}.$$

Lemma 2. *Any single-agent $\mathcal{L}(\Sigma)$ epistemic state (M, w) is bisimilar to its $\mathcal{M}_{(\Sigma, D)}$ substate $(M^\star, w^\star)$.*

Proof. Define $Z_f \subseteq W \times W^\star$ by $vZ_f v'$ iff $R_d wv$ and $\forall x \in \mathtt{C} \cup \mathtt{R}, I(x, v) = I^\star(x, v')$, where f is the identity map. Then Z_f is a bisimulation and it links w to $w^\star$:
Atoms: If $vZ_f v'$ then $I(x, v) = I^\star(x, v'), \forall x \in \mathtt{C} \cup \mathtt{R}$, which, since M and $M^\star$ share domain D, entails v and v' are elementary equivalent. As D is finite, by Lemma 1, $v \simeq_f v'$. **Forth**: Assume that $vZ_f v'$ and $R_d vu$. By definition of $M^\star$, there is a $u' \in W^\star$ s.t. $\forall x \in \mathtt{C} \cup \mathtt{R}, I(x, u) = I^\star(x, u')$. By definition of Z_f, we have $R_d wv$ and so, by transitivity, that $R_d wu$. Hence $uZ_f u'$. Finally, as $R_d^\star = W^\star \times W^\star$, we have $R_d^\star v'u'$. **Back**: Assume that $vZ_f v'$ and $R_d^\star v'u'$. By definition of $M^\star$, there is a $u \in W$ such that $R_d wu$ and $I(x, u) = I^\star(x, u'), \forall x \in \mathtt{C} \cup \mathtt{R}$. Hence, $uZ_f u'$. As $vZ_f v'$, by definition, we have $R_d wv$. As also $R_d wu$, we have $R_d vu$, by Euclideaness. **Link**: Finally, as $I(x, w) = I^\star(x, w^\star), \forall x \in \mathtt{C} \cup \mathtt{R}$, and $R_d ww$ by reflexivity, it follows that $wZ_f w^\star$. $\square$

Lemma 3. *There are only finitely many non-bisimilar single-agent $\mathcal{L}(\Sigma)$ epistemic states with underlying finite domain D.*

Proof. By Lemma 2, any single-agent $\mathcal{L}(\Sigma)$ epistemic state (M, w) with domain D is bisimilar to a $\mathcal{M}_{(\Sigma, D)}$ substate. Hence, up to bisimulation, the set of single-agent $\mathcal{L}(\Sigma)$ epistemic states is representable by a set of $\mathcal{M}_{(\Sigma, D)}$ substates. Since both Σ and D are finite, there are only finitely many first-order interpretations of Σ over D—i.e., $\mathcal{W}$ is finite. Hence $\mathcal{M}_{(\Sigma, D)}$ has finitely many substates. $\square$

Theorem 1. $\mathsf{PlanEx}(1)$ *(single-agent planning) is decidable.*

Proof. Let $P = (s_0, \mathcal{A}, \varphi_G)$ be a single-agent planning task. By Proposition 1 and Lemma 2, any epistemic state and its $M_{(\Sigma, D)}$ substate are bisimilar and hence agree on φ_G. By Proposition 4, bisimilarity is preserved under product update. Thus, to check if there is a plan, we do a breadth-first search of the state space, replacing each visited epistemic state by its $M_{(\Sigma, D)}$ substate. By Lemma 3, this state space is finite. $\square$

4.2 Multi-Agent Planning

Concerning multi-agent planning, the plan existence problem is undecidable for unrestricted FODEL planning. This follows directly from the problem being undecidable for DEL planning and FODEL containing DEL as a special case: any DEL language can be simulated with a signature Σ consisting only of 0-ary predicates. Aucher and Bolander [2013] show that DEL planning is undecidable already for two agents.

Proposition 6. *For $k \geq 2$, $\mathsf{PlanEx}(k)$ is undecidable.*

Reasonably expressive DEL fragments are decidable, such as the fragment in which all actions have propositional preconditions. We show that this result generalizes to FODEL. Our result is inspired by [Yu et al., 2013], which shows decidability for DEL. As in the single-agent case, the proof exploits the fact that the state space in such case can be shown to be finite (up to bisimulation). We fix some lemmas.

Lemma 4. *There are finitely many non-bisimilar $\mathcal{L}(\Sigma)$ states with finite domain D and exactly n worlds.*

Proof. Let $|W| = n$. As Σ and D are finite, the set $\mathcal{I}$ of interpretations definable for W, Σ and D is finite. As $D_{\mathtt{agt}} \subseteq D$ is finite, also the set $\mathcal{R}$ of definable accessibility relations for $D_{\mathtt{agt}}$ and W is finite. Hence the set of states $\{(M, w) \mid D^M = D, W^M = W, R^M \in \mathcal{R}, I^M \in \mathcal{I}, w \in W\}$ is finite. As the choice of W is irrelevant for bisimilarity, there are thus only finitely many non-bisimilar states with n worlds. $\square$

Yu *et al.* [2013] showed that the number of $\underline{\leftrightarrow}_n$ equivalence classes of Kripke model has a fixed finite bound that depends on n, generalizing a result from Visser [1996]. We show a similar result for first-order models.

Lemma 5. *Let M be an $\mathcal{L}(\Sigma)$ model with finite domain D. Let $|D| = m$ and $|D_{\mathtt{agt}}| = \ell$. Denote by $F_n(M)$ the number of $\underline{\leftrightarrow}_n$ equivalence classes of worlds of M. Let $f(0) := m!$ and $f(i + 1) := m! 2^{f(i) + \ell}$ for $i \geq 0$. Then $F_n(M) \leq f(n)$.*

Proof. By induction on k. Let $M = (D, W, R, I)$. Base case: $k = 0$. If $w \simeq_f v$, then $M, w \underline{\leftrightarrow}_0 M, v$. As there are $m!$ bijections from D into D, there are at most $m!$ non-isomorphic worlds in M and thus at most $m!$ $\underline{\leftrightarrow}_0$ equivalence classes. Hence $F_0(M) \leq f(0)$. Inductive step: Assume that the statement holds for k. We prove it for $k + 1$. For any $w \in W$ and $d \in D_{\mathtt{agt}}$, let $R_d[w] := \{[w']_{\underline{\leftrightarrow}_k} \mid R_d ww'\}$. Then, for any $w_1, w_2 \in W$, if $w_1 \simeq_f w_2$ and $R_d[w_1] = R_d[w_2]$ for all $d \in D_{\mathtt{agt}}$, then $(M, w_1) \underline{\leftrightarrow}_{k+1} (M, w_2)$. For any $d \in D_{\mathtt{agt}}$, there are at most $2^{F_k(M)}$ distinct possibilities for $R_d[w]$. Hence $F_{k+1}(M) \leq m! (2^{F_k(M)})^\ell \leq m! (2^{f(k)})^\ell = m! 2^{f(k) + \ell} = f(k + 1)$. $\square$

Theorem 2. *If all actions have non-modal preconditions, then $\mathsf{PlanEx}(k)$ is decidable, for $k \geq 1$.*

Proof. Let $P = (s_0, \mathcal{A}, \varphi_G)$ be an epistemic planning task with $md(\varphi_G) = n$. By Proposition 3, any epistemic state and its n-bisimulation contraction are n-bisimilar and thus agree on φ_G. By Proposition 5, product update preserves n-bisimilarity. To check for a plan, we do a breath-first search of the state space, replacing each visited state by its n-bisimulation contraction. By Lemma 5, any n-bisimulation contracted model has at most $F_n(M) \in \mathbb{N}$ worlds and by Lemma 4, the number of non-bisimilar states with at most $F_n(M)$ worlds is finite. $\square$

5 Final Remarks

FODEL combines the higher-order expressiveness of DEL with the first-order abstraction of languages such as PDDL to compactly describe epistemic planning domains. We have introduced *bisimulations* for FODEL models and exploited them to show *decidability* for single-agent planning and multi-agent planning with non-modal preconditions. Future work could study other decidable fragments, the complexity of FODEL planning, or FODEL planning algorithms.

Acknowledgements

References

[Achen *et al.*, 2020] Andreas Achen, Andrés Occhipinti Liberman, and Rasmus K. Rendsvig. Dynamic Term-Modal Logics for First-Order Epistemic Planning. *Artificial Intelligence*, forthcoming, 2020.

[Aucher and Bolander, 2013] Guillaume Aucher and Thomas Bolander. Undecidability in Epistemic Planning. In *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.

[Baltag *et al.*, 1998] Alexandru Baltag, Lawrence S. Moss, and Sławomir Solecki. The Logic of Public Announcements, Common Knowledge, and Private Suspicions (extended abstract). In *TARK '98: Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 43–56. Morgan Kaufmann Publishers, 1998.

[Baral *et al.*, 2017] Chitta Baral, Thomas Bolander, Hans van Ditmarsch, and Sheila McIlrath. Epistemic Planning (Report from Dagstuhl Seminar 17231). *Dagstuhl Reports*, 7(6):1–47, 2017.

[van Benthem *et al.*, 2006] Johan van Benthem, Jan van Eijck, and Barteld P. Kooi. Logics of Communication and Change. *Information and Computation*, 204(11):1620–1662, 2006.

[van Benthem, 2010a] Johan van Benthem. Frame Correspondences in Modal Predicate Logic. In S. Feferman and W. Sieg, editors, *Proofs, Categories and Computations. Essays in Honor of Grigori Mints*. College Publications, 2010.

[van Benthem, 2010b] Johan van Benthem. *Modal Logic for Open Minds*. CSLI, 2010.

[Blackburn and van Benthem, 2007] Patrick Blackburn and Johan van Benthem. Modal Logic: A Semantic Perspective. In *Studies in Logic and Practical Reasoning*, volume 3, pages 1–84. Elsevier, 2007.

[Blackburn *et al.*, 2001] Patrick Blackburn, Maarten De Rijke, and Yde Venema. *Modal Logic*. Cambridge University Press, 2001.

[Bolander and Andersen, 2011] Thomas Bolander and Mikkel Birkegaard Andersen. Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics*, 21(1):9–34, 2011.

[Bolander, 2014] Thomas Bolander. Seeing is Believing: Formalising False-Belief Tasks in Dynamic Epistemic Logic. In Andreas Herzig and Emiliano Lorini, editors, *European Conference on Social Intelligence (ECSI 2014)*, volume 1283 of *CEUR Workshop Proceedings*, pages 87–107, 2014.

[Bolander, 2017] Thomas Bolander. A Gentle Introduction to Epistemic Planning: The DEL Approach. In *Proceedings of Methods for Modalities 2017 (M4M9)*, volume 243 of *Electronic Proceedings of Theoretical Computer Science*, pages 1–22, 2017.

[van Ditmarsch *et al.*, 2005] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld P. Kooi. Dynamic Epistemic Logic with Assignment. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '05)*, pages 141–148. ACM, 2005.

[Ghallab *et al.*, 2004] Malik Ghallab, Dana Nau, and Paolo Traverso. *Automated Planning: Theory and Practice*. Morgan Kaufmann, 2004.

[Kripke, 1962] Saul A. Kripke. The Undecidability of Monadic Modal Quantification Theory. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 8:113–116, 1962.

[Occhipinti Liberman and Rendsvig, 2019] Andrés Occhipinti Liberman and Rasmus K. Rendsvig. Dynamic Term-Modal Logic for Epistemic Social Network Dynamics. In Patrick Blackburn, Emiliano Lorini, and Meiyun Guo, editors, *Logic, Rationality, and Interaction (LORI 2019)*, volume 11813 of *Lecture Notes in Computer Science*, pages 168–182. Springer, 2019.

[Rothmaler, 2018] Philipp Rothmaler. *Introduction to Model Theory*. CRC Press, 2018.

[Visser, 1996] Albert Visser. Uniform Interpolation and Layered Bisimulation. In Petr Hájek, editor, *Gödel '96: Logical Foundations of Mathematics, Computer Science and Physics – Kurt Gödel's Legacy*, volume 6 of *Lecture Notes in Logic*, pages 139–164. Cambridge University Press, 1996.

[Yu *et al.*, 2013] Quan Yu, Ximing Wen, and Yongmei Liu. Multi-Agent Epistemic Explanatory Diagnosis via Reasoning about Actions. In *Twenty-Third International Joint Conference on Artificial Intelligence*, pages 1183–1190, 2013.

[Zoghifard and Pourmahdian, 2018] Reihane Zoghifard and Pourmahdian Pourmahdian. First-Order Modal Logic: Frame Definability and a Lindström Theorem. *Studia Logica*, 106(4):699–720, 2018.