# On the Role of Postconditions in Dynamic First-Order Epistemic Logic

Côme Neyrand[1] and Sophie Pinchinat[2]

[1] IRISA/Univ Rennes come.neyrand@ens-rennes.fr
[2] IRISA/Univ Rennes sophie.pinchinat@irisa.fr

**Abstract.** Dynamic Epistemic Logic (DEL) is a logic that models information change in a multi-agents setting through the use of action models with pre and post-conditions. In a recent work, an extension of DEL to first-order epistemic logic (DFOEL) was introduced with a proof of the decidability of Epistemic Planning in this setting for non-modal pre and post-conditions under finite-domain hypothesis.
In this original contribution, we exhibit the role post-conditions have in DFOEL by showing that Epistemic Planning in this setting with possibly infinite first-order domain is undecidable as soon as even non-modal post-conditions contain first-order quantifiers, while Epistemic Planning becomes decidable when post-conditions are quantifier-free. The latter result is non-trivial and makes an extensive use of automatic structures.

## 1 Introduction

First-order modal logic has been around since the first introduction of modal logic. Modal logic (see [2]) was introduced as an extension of propositional logic, its semantics applies to Kripke models i.e., different valuations in modal relations to one another. These modal relations can be of different sorts creating a wide range of logics that are direct applications of modal logic. For example, temporal logic and epistemic logic are modal logics in which modal relations are defined as relations of time and knowledge respectively. First-order modal logic is an extension of modal logic in which valuations are replaced by first-order structures, each modal logic can be extended to its first-order counterpart (see [2, Chapter 9]).
A remarkable application of modal logic, Dynamic Epistemic Logic (DEL) [16], provides a logic that describe changes on an epistemic model through the use of action models and update. The links between DEL and Epistemic Temporal Logic (ETL) have been studied in [12]. Under some hypothesis on DEL action models, the underlying ETL model is regular, allowing to decide central problems in artificial intelligence: a typical problem is Epistemic Planning – an automated planning problem in the DEL setting introduced by [4] – which has been thoroughly studied (see [5] for a survey). Recently an extension of Epistemic Planning in the setting of Dynamic First-Order Epistemic Logic (DFOEL) has been introduced in [13] – following the pioneer approach of [11] – where first-order

logic is used as a means to compactly specify an essentially propositional input.

In [14], the authors prove that so-called non-modal epistemic planning in DFOEL remains decidable provided that the first-order domain is finite. Even though one can foresee a full propositional encoding of the entire problem hence the decidability, the authors of [14] provide an elegant proof based on bisimulations. However this bisimulation approach does not apply when we relax the finiteness assumption on first-order domains.

In this paper we investigate Epistemic Planning in the setting of DFOEL with possibly infinite domain structures. Our aim is to enlarge the decidability landscape for Epistemic Planning, we therefore discard the hopeless case of events with modal operators by focusing on pure first-order events. Surprisingly, our contribution shows that even though preconditions are non-modal, decidability of Epistemic Planning is sensitive to the nature of post-conditions.

More precisely, we consider Epistemic Planning where the input structures are finally representable. As a natural approach, we require the input epistemic model to be automatic. Regarding the action models, we allow arbitrary first-order interpretations thus strictly extending the framework of [13]. On this basis, we show that Epistemic Planning with arbitrary first-order post-conditions (EPP1) is undecidable while Epistemic Planning with quantifier-free postconditions (EPP0) is decidable. Our decidability proof for EPP0 is non-trivial because the infinite domain structures generated along histories are in general infinitely many.

The paper is organized as follows. In Section 2, we define the full framework of DFOEL. In Section 3 we introduce the Epistemic Planning problem in this framework and establish our results: undecidability of EPP1 in Section 3.4 and decidability of EPP0 in Section 3.5. We conclude on these achievements in Section 4.

## 2 Dynamic First-Order Epistemic Logic

In this section we describe our proposal for DFOEL, inspired from [13] but that offers a wider expressiveness in both the epistemic and the actions models: first, we allow one to consider infinite (but still finitely presentable) first-order structures, and second, we relax the action model post-conditions, i.e., the predicate updates, as arbitrary *first-order interpretations* (in the sense of model theory, see [8]) but where, for semantical reasons, the domain remains unchanged.

We will see that this extra expressiveness yields an undecidable epistemic planing problem. We will then better control the expressiveness to retrieve decidability, still in a setting that strictly extends the results from [13] for allowing quantifier-free predicate updates, that in general may involve an infinite set of tuples.

For the rest of this paper, we fix $P$ a finite set of predicate symbols, $X$ a countably infinite set of variables and `agt` a finite set of agents.

### 2.1 Preliminaries on First-Order Epistemic Logic

We restrict our definition to pure relational first-order structures (i.e., no functions in the signature)[3].

**Definition 1.** *The* language *of First-Order Epistemic Logic (FOEL) $\mathcal{L}$ is given by the following syntax:*

$$\mathcal{L} \ni \varphi, \psi ::= p(x_1, \ldots, x_n) \mid \neg\varphi \mid \varphi \wedge \psi \mid \forall x \varphi \mid K_a \varphi$$

*where $x, x_1, ..., x_n \in X$, $a \in \mathtt{agt}$ and $p \in P$. An* atomic formula *is a formula of the form $p(x_1, \ldots, x_n)$. We denote by $\mathcal{L}_1$ the $K_a$-free sub-language of $\mathcal{L}$, and by $\mathcal{L}_0$ the $K_a$-free and quantifier-free sub-language of $\mathcal{L}$.*

In epistemic logic, models are represented by Kripke models with possible worlds related through epistemic relations: intuitively, two worlds are in epistemic relationship for an agent $a$ if $a$ cannot distinguish between the two. In the setting of DFOEL, an entire first-order structure is attached to each possible word.

**Definition 2.** *A* first-order epistemic model *(or simply* epistemic model*) over domain $D$ is a structure $\mathcal{M} = (W, (R_a)_{a \in \mathtt{agt}}, (\mathcal{I}_w)_{w \in W})$ where:*

1. *$W$ is a non-empty set of possible worlds;*
2. *For each $a \in \mathtt{agt}$, $R_a \subseteq W \times W$ is an accessibility relation;*
3. *For every $w \in W$, $\mathcal{I}_w = (D, (p^w)_{p \in P})$ is a $P$-structure associated to $w$.*

*Note that the structures $\mathcal{I}_w$ share the same domain $D$, but that the predicates may have different interpretations in different worlds.*

**Definition 3.** *Let $\mathcal{M}$ be an epistemic model over domain $D$. A* valuation *is a mapping $v : X \to D$, and a $x$-*variant *$v_x$ of $v$ is a valuation s.t. $v_x(y) = v(y)$ for all $y \in X \backslash \{x\}$.*

**Definition 4.** *Let $\mathcal{M}$ be a model and $v$ be a valuation. The* satisfaction *relation between an epistemic model $\mathcal{M}$ and a formula of $\mathcal{L}$ is given inductively by :*

- *$\mathcal{M}, w \models_v p(x_1, \ldots, x_n)$ iff $(v(x_1), ..., v(x_n)) \in p^w$ for all $p \in P$.*
- *$\mathcal{M}, w \models_v \neg\varphi$ iff not $\mathcal{M}, w \models_v \varphi$.*
- *$\mathcal{M}, w \models_v \varphi \wedge \psi$ iff $\mathcal{M}, w \models_v \varphi$ and $\mathcal{M}, w \models_v \psi$.*
- *$\mathcal{M}, w \models_v \forall x \varphi$ iff $\mathcal{M}, w \models_{v_x} \varphi$ for every $x$-variant $v_x$ of $v$.*
- *$\mathcal{M}, w \models_v K_a \varphi$ iff $\mathcal{M}, w' \models_v \varphi$ for all $w' \in R_a(w)$.*

*Note that if $\varphi \in \mathcal{L}_1$ i.e., $\varphi$ is modal-free, then $\mathcal{M}, w \models_v \varphi$ iff $\mathcal{I}_w \models_v \varphi$. We write $\mathcal{M}, w \models \varphi$ whenever for any $v$, $\mathcal{M}, w \models_v \varphi$; in particular if $\varphi$ has no free variables.*

It is clear that classical propositional epistemic logic is a fragment of FOEL where predicate symbols have all arity 0.

[3] This is no loss of expressiveness since functions and constants can be modeled by predicates through their graph.

## 2.2 Dynamic First-Order Epistemic Logic

We now enrich the setting of FOEL with *action models* that provide the dynamics, via the *update product*, similarly to the approach in DEL. An action is a Kripke model whose elements are *events* with their *precondition* and their *postconditions*.

**Definition 5.** *An* action model *is a tuple* $\mathcal{E} = (E, (Q_a)_{a\in \mathtt{agt}}, pre, post)$ *where:*
1. *$E$ is a non-empty finite set of possible events.*
2. *For $a \in \mathtt{agt}$, $Q_a \subseteq E \times E$ is an accessibility relation.*
3. *$pre : E \to \mathcal{L}$ assigns to each $e \in E$ a precondition formula without free-variables.*
4. *Let $\{z_1, z_2, ...\}$ be a set of variables disjoint from $X$.*
   *$post : E \to (P \to \mathcal{L})$ assigns to each $e \in E$ a postcondition for each predicate s.t. $post(e)(p)$ has its free-variables in $\{z_1, ..., z_{kp}\}$.*

The update of an epistemic state with an action model filters the worlds that verify the preconditions and updates their structure through the postconditions: the domain is unchanged, but predicates might be updated by an arbitrary *first-order interpretation*.

**Definition 6.** *Let* $\mathcal{M} = (W, (R_a)_{a\in \mathtt{agt}}, (\mathcal{I}_w)_{w\in W})$ *and* $\mathcal{E} = (E, (Q_a)_{a\in \mathtt{agt}}, pre, post)$ *be given. The* product update *of* $\mathcal{M}$ *and* $\mathcal{E}$ *yields the model*

$$\mathcal{M} \otimes \mathcal{E} = (W', ({R'}_a)_{a\in \mathtt{agt}}, (\mathcal{I}_w)_{w\in W'})$$

*where :*
1. *$W' = \{(w, e) \in W \times E : \mathcal{M}, w \models pre(e)\}$. Denote by $we$ the world $(w, e)$.*
2. *For $a \in \mathtt{agt}$, $we{R'}_a w'e'$ iff $wR_a w'$ and $eQ_a e'$.*
3. *For $we \in W'$, $\mathcal{I}_{we} = (D, (p^{we})_{p\in P})$ where for $p \in P$,*
   *$p^{we} = \{(d_1, ..., d_n) \mid \mathcal{M}, w \models_{[z_1 \mapsto d_1,...,z_n \mapsto d_n]} post(e)(p)(z_1, ..., z_k)\}$*

A post-condition can be for example $post(e)(p)(z_1, z_2) = p(z_1, z_2) \vee \exists x(p(z_1, x) \wedge p(x, z_2))$ which computes the transitive closure of $p$ after multiple updates.

## 2.3 The infinite epistemic model of histories

We aim at capturing the single infinite epistemic model comprised of all the updates. In a way similar to the DEL structure introduced in [1, 5] for proposition DEL, it consist in putting together the iterated updates. Given $\mathcal{M}$ be an epistemic model and $\mathcal{E}$ be an action model, we define the family of updates $(\mathcal{ME}^n)_{n\in\mathbb{N}}$ defined by: $\mathcal{ME}^0 = \mathcal{M}$ and $\mathcal{ME}^{n+1} = \mathcal{ME}^n \otimes \mathcal{E}$. We denote by $\mathcal{ME}^n = (W_n, (R^n_a)_{a\in \mathtt{agt}}, (\mathcal{I}_h)_{h\in W_n})$ the $n$-th update.
Notice that elements $h \in W_n$ are of the form $w'e_1...e_n$, where $w' \in W$ and $e_1, \dots, e_n \in E$. We call such an element a *history starting in $w'$*.

**Definition 7 (The epistemic model of histories).** *The* epistemic model of histories *is* $\mathcal{ME}^* = (H, (R_a)_{a\in \mathtt{agt}}, (\mathcal{I}_h)_{h\in H})$*, where* $H := \bigcup_{n\in\mathbb{N}} W_n$ *is the set of* histories *and where* $R_a := \bigcup_{n\in\mathbb{N}} R^n_a$ *is the accessibility relation over all the histories for agent for $a \in \mathtt{agt}$.*
*We write $H_w$ for the set of histories starting in world $w \in W$.*

# 3 Epistemic Planning Problems in First-order DEL

Now that the epistemic model $\mathcal{ME}^*$ of histories is defined, one can address decision problems, as long as this structure can be finitely represented as the input of an algorithm. As $\mathcal{ME}^*$ is fully defined by $\mathcal{M}$ and $\mathcal{E}$, it amounts to defining both in a finite way. Since $\mathcal{E}$ is finite, it remains to provide a way to describe the first-order epistemic model $\mathcal{M}$. A standard approach to finitely presenting structures is to resort to finite-state automata, leading to the notion of *automatic structure*. It is important to remark that the first-order theory of any automatic structure is decidable, thanks to automata constructions that mimic the logical operations. We refer the reader to [15] for a survey.
In the following section we provide the material to define the class of first-order epistemic models with finitely many worlds but whose first-order interpretations in worlds are automatic structures. Then we introduce the epistemic planning problem in this setting and discuss its decidability.

## 3.1 Automatic presentations of epistemic models

Automatic structures are first-order structures with no function symbols, such that its domain and predicates are regular, that is recognized by finite-state automata. *Automatically presentable* structures are relational structures isomorphic to an automatic structure (see [15] for an exhaustive survey).

**Definition 8.** *An* automatic presentation *over the set of predicates $P$ is a finite tuple $A = (\mathcal{A}_D, (\mathcal{A}_p)_{p \in P})$ of finite-state automata where*
- *$\mathcal{A}_D$ is an automaton over alphabet $\Sigma$, and*
- *for $p \in P$ of arity $k$, $\mathcal{A}_p$ is an automaton over alphabet $(\Sigma \cup \{\square\})^k$.*

*An automatic presentation $A = (\mathcal{A}_D, (\mathcal{A}_p)_{p \in P})$ denotes a first-order structure over signature $P$ defined by $\mathcal{S}_A = \langle L(\mathcal{A}_D), L(\mathcal{A}_p)_{p \in P} \rangle$.*
*A first-order structure $\mathcal{S} = (S, (p^{\mathcal{S}})_{p \in P})$ over signature $P$ is* automatic *if there exists a bijection isomorphism $\texttt{enc} : S \to L(\mathcal{A}_D)$ for some automatic presentation $A = (\mathcal{A}_D, (\mathcal{A}_p)_{p \in P})$ that is an isomorphism between $\mathcal{S}$ and $\mathcal{S}_A$.*

Automatic structures have a smooth connection with first-order logic:

**Theorem 1 ([3]).** *The first-order theory of an automatic structure is decidable.*

In our setting, we allow the domain of the (first-order) epistemic models to be infinite as long as they are automatic, yielding the natural notion of *automatic epistemic models.*

**Definition 9 (Automatic epistemic models).** *An epistemic model $\mathcal{M} = (W, (R_a)_{a \in agt}, (\mathcal{I}_w)_{w \in W})$ is* automatic *if the set of worlds $W$ is finite and for each $w \in W$, the structure $\mathcal{I}_w$ is automatic.*

It is easy to see that the predicates updates $post(e)(p)$ for a non-modal action model are mere *first-order interpretations* which by [3, Proposition 5.2] preserve automaticity.

**Proposition 1** *Let $\mathcal{E} = (E, (Q_a)_{a \in agt}, pre, post)$ be a non-modal action model. If $\mathcal{M}$ is an automatic epistemic model, then $\mathcal{M} \otimes \mathcal{E}$ is also an automatic epistemic model.*

### 3.2 The Epistemic Planning Problem

As originally defined in [4], an instance of the epistemic planning problem is composed of an automatic epistemic model, a distinguished world in this model, an action model and a first-order epistemic formula called the *goal formula*. The problem is to decide whether or not there exists an executable sequence of events from the distinguished world so that the goal formula holds. Restated in the setting we have developed so far, the epistemic planning problem is defined as follows.

**Definition 10 (The Epistemic Planning Problem (EPP)).**
**Input:** *an automatic epistemic model $\mathcal{M}$, a distinguished world $w$ in $\mathcal{M}$, an action model $\mathcal{E}$ and a first-order epistemic formula $\gamma$;*
**Output:** *is there a history $h \in H_w$ such that $\mathcal{M}\mathcal{E}^*, h \models \gamma$?*

Epistemic planning has been widely investigated in the literature [5, 4, 6, 7, 1, 13, 14]. The first family of contributions [5, 4, 6, 7, 1] consider propositional epistemic and actions models. It is clear that the propositional variant of the epistemic problem is a sub-problem of EPP, since propositional logic can be embedded in first-order logic. In this propositional setting, it is well-known that the problem is undecidable if one allows the event pre-conditions to be modal formulas ([5]). As a corollary, we can state the following:

**Theorem 2.** *EPP is undecidable.*

Because it is known that modal operators easily navigate between pre and post-conditions, we consider a restricted variant of EPP, written EPP1, where pre-conditions and post-conditions are non-modal i.e., each even $e$, formulas $pre(e)$ and $post(e)(p)$ belong to $\mathcal{L}_1$.
We show that the problem EPP1 is undecidable (Theorem 3) which shows that the decidability result of [13] fails when the first-order domains can be infinite.Nevertheless there is a place for a decidable restriction of EPP1, written EPP0, where post-conditions of the input action models are non-modal and quantifier-free i.e., all formulas $post(e)(p)$ belong to $\mathcal{L}_0$ (Theorem 4). Remark that the decidability of Epistemic Planning in the DEL setting (propositional action models) is corollary of the decidability of EPP0.

As a preliminary step that will be useful for our result, we approach EPP under the light of pure first-order logic in a structure derived from the epistemic model of histories.

### 3.3 The first-order structure of the epistemic model of histories

Given an epistemic model $\mathcal{ME}^* = (H, (R_a)_{a\in\mathtt{agt}}, (\mathcal{I}_h)_{h\in H})$, we consider the universe, i.e., the domain, $U$ comprised of every history $h \in H$, but also of countably many disjoint copies of $D$, a copy $D_h$, for each $h \in H$, that denotes the elements of the structure $\mathcal{I}_h$. We equip $U$ with several relations in order to obtain a structure $\mathbb{H}$ (Definition 11), and that yields a first-order language such that the problem EPP reduces into the model-checking problem over $\mathbb{H}$ against formulas in this language (Proposition 3).
Theorem 4 is then an immediate corollary of Proposition 3, Theorem 5 and Theorem 1. We now accurately formalize all this.
We reshape each predicate $p \in P$ into predicate $\mathtt{p}^*$ of arity $ar(p)+1$ and introduce a finite number of extra predicates. We list all these predicates below to form signature $\tau = ((\mathtt{R}_a)_{a\in\mathtt{agt}}, (p^*)_{p\in P}, (\mathtt{from}_w)_{w\in W}, \mathtt{Dom})$, with their interpretation in $\mathbb{H}$.

- predicate $\mathtt{p}^*$ of arity $ar(p)+1$, for each $p \in P$, where ${\mathtt{p}^*}^{\mathbb{H}}$ is the set of tuples $\{(h, d_1, \ldots, d_k) \mid (d_1, \ldots, d_k) \in p^h]\}$;
- predicate $\mathtt{R}_a$ of arity 2, for each agent $a \in \mathtt{agt}$, where ${\mathtt{R}_a}^{\mathbb{H}}$ – its interpretation in $\mathbb{H}$ – is the pairs of histories related by $R_a$ in $\mathcal{ME}^*$;
- predicate $\mathtt{from}_w$ of arity 1, for each $w \in W$, where ${\mathtt{from}_w}^{\mathbb{H}}$ is the subset $H_w$ of histories that start from $w$;
- predicate $\mathtt{Dom}$ of arity 2, where $\mathtt{Dom}^{\mathbb{H}}$ relates a history with every element of domain $D_h$.

We can now describe the first-order structure we need to handle in our proof.

**Definition 11 (The structure $\mathbb{H}$).** *Let $\mathcal{ME}^*$ be an epistemic model of histories, with $D$ the domain shared by all the histories. For each $h \in H$, we let $D_h$ be a disjoint copy of $D$ and we write $\mathtt{copy}_h : D \to D_h$ for the natural bijection, and we define the set $U := H \cup \biguplus_{h\in H} D_h$. Then, the* first-order structure of $\mathcal{ME}^*$ *is the $\tau$-structure defined by:*

$$\mathbb{H} = (U, (R_a)_{a\in agt}, (\bigcup_{h\in H} (\{h\} \times p^h))_{p\in P}, (H_w)_{w\in W}, \bigcup_{h\in H} (\{h\} \times D_h))$$

The structure $\mathcal{ME}^*$ equipped with first-order logic over signature $\tau = ((\mathtt{R}_a)_{a\in\mathtt{agt}}, (p^*)_{p\in P}, (\mathtt{from}_w)_{w\in W}, \mathtt{Dom})$ is at least as expressive as the epistemic model of histories, as stated by Proposition 2. We first recall the standard translation of FOEL into first-order logic as done in [2].

**Definition 12.** *The standard translation $ST_y$ of $\mathcal{L}$ into the first-order logic over signature $\tau$ is inductively defined by:*

- $ST_y(p(x_1, \ldots, x_n)) := \mathtt{p}^*(y, x_1, \ldots, x_n) \wedge \bigwedge_{i=1}^n \mathtt{Dom}(y, x_i)$
- $ST_y(\neg\varphi) := \neg ST_y(\varphi)$
- $ST_y(\varphi \wedge \psi) := ST_y(\varphi) \wedge ST_y(\psi)$
- $ST_y(\forall x\varphi(x)) := \forall x(\mathtt{Dom}(y, x) \to ST_y(\varphi(x)))$
- $ST_y(K_a\varphi) := \forall y'(\mathtt{R}_a(y, y') \to ST_{y'}(\varphi))$

**Proposition 2** *For any valuation $v : X \to D$ and any history $h \in H$, we let $v_h : X \to D_h$ be defined by $v_h(x) = \mathtt{copy}_h(v(x))$.*

$$\mathcal{ME}^*, h \models_v \varphi \textit{ iff } \mathbb{H} \models_{v_h[y \mapsto h]} ST_y(\varphi)$$

*Proof.* We proceed by induction over $\varphi$:

- $\mathcal{ME}^*, h \models_v p(x_1, \ldots, x_n)$ iff $(v(x_1), ..., v(x_n)) \in p^h$
  iff (by definition of $\mathtt{p}^{*\mathbb{H}}$) $(h, v_h(x_1), ..., v_h(x_n)) \in \mathtt{p}^{*\mathbb{H}}$
  iff $\mathbb{H} \models_{v_h[y \mapsto h]} \mathtt{p}^*(y, x_1, ..., x_n)$
  iff $\mathbb{H} \models_{v_h[y \mapsto h]} \mathtt{p}^*(y, x_1, ..., x_n) \wedge \bigwedge_{i=1}^n \mathtt{Dom}(y, x_i)$, and this latter formula is precisely $ST_y(p(x_1, \ldots, x_n))$;
- the cases for formulas of the form $\neg\varphi$ and $\varphi \wedge \psi$ is smooth;
- $\mathcal{ME}^*, h \models_v \forall x \varphi$
  iff $\mathcal{M}, h \models_{v_x} \varphi$ for every $x$-variant $v_x$ of $v$
  iff (by induction) $\mathbb{H} \models_{(v_x)_h[y \mapsto h]} ST_y(\varphi)$ for every $x$-variant $(v_x)_h$ of $v_h$;
- $\mathcal{ME}^*, h \models_v K_a\varphi$ iff for all $h' \in R_a(h)$, $\mathcal{M}, h' \models_v \varphi$
  iff for all $h' \in R_a(h)$, $\mathbb{H} \models_{v_{h'}[y' \mapsto h']} ST_{y'}(\varphi)$ (by induction)
  iff $\mathbb{H} \models_{v_h[y \mapsto h]} \forall y'(\mathtt{R}_a(y, y') \to ST_{y'}(\varphi))$ (since $D_h = D_{h'}$)
  iff $\mathbb{H} \models_{v_h[y \mapsto h]} ST_y(K_a\varphi)$ (by definition of $ST_y$).

An immediate corollary of Proposition 2 is a reduction of EPP into the model-checking problem over $\mathbb{H}$ against a first-order logic:

**Proposition 3** *There exists $h \in H_w$ s.t. $\mathcal{ME}^*, h \models \gamma$ if, and only if, $\mathbb{H} \models \exists y(ST_y(\gamma) \wedge \mathit{from}_w(y))$.*

### 3.4 Undecidability of EPP1

We can demonstrate that with a $post(e)(p) \in \mathcal{L}_1$, we can capture the transitive closure of a graph which allows us to reduce the emptiness problem on Turing machines (TME) to EPP1. The problem TME is known to be undecidable [9, Theorem 9.10] and is defined as follows.

**Input:** a Turing machine $M$;
**Output:** do we have $L(M) = \emptyset$?

**Theorem 3.** *The Epistemic Planning Problem EPP1 is undecidable.*

The proof of Theorem 3 can be found in Appendix A. It relies on the fact that the configuration graph of a Turing machine is automatically presentable [10, Lemma 5.1] and that by updating the successor predicate in this graph by transitivity one can approximate its transitive closure.

### 3.5 Decidability of EPP0

To establish the decidability of EPP0, we take inspiration from [5], where the decidability of epistemic planning in DEL under non-modal assumptions (a sub-problem of EPP0) is developed. The proof relies on the fact

that the structure that we have named here the epistemic model of histories (structure $\mathcal{ME}^*$) is automatic and the fact that epistemic planning reduces to model-checking first-order formula over $\mathcal{ME}^*$.
In a nutshell, we show in this section that the first-order structure $\mathbb{H}$ of the epistemic model of histories $\mathcal{ME}^*$ is automatic if one restricts to inputs allowed in the problem EPP0. Since by Proposition 3, we have already established that EPP reduces to the model checking over $\mathbb{H}$ against first-order logic and because first-order logic is decidable on automatic structures (Theorem 1), we can conclude the following.

**Theorem 4.** *EPP0 is decidable.*

The rest of this section is dedicated to the proof of the following Theorem 5.

**Theorem 5.** *Let $\mathcal{M}$ be an automatic epistemic model and $\mathcal{E}$ be such that all pre and post-conditions are non-modal, and post-conditions are quantifier-free.*
*Then the first-order structure $\mathbb{H}$ of the epistemic model of histories $\mathcal{ME}^*$ is automatic.*

As preconditions and postconditions are non-modal, the update of the predicate interpretations at some history $h \in H$ by an event $e$ only depends on $\mathcal{I}_h$ and $e$. We can thus keep track of the interpretation along $we_1...e_n$ after the trigger of each event $e_i$ by remembering the current interpretation. Now, since postonditions are quantifer-free, only a finite number of possible interpretations of predicates are generated by the successive updates (Proposition 4). This allows us to define an automaton that recognizes the set of histories sharing a common predicate interpretation (Proposition 5). We then define an encoding function `enc` for the entire domain of the structure $\mathbb{H}$ and show that it yields regular languages for $\texttt{enc}(U)$ (Lemma 1) but also for $\texttt{enc}(R_a^{\mathbb{H}})$, $\texttt{enc}(p^{*\mathbb{H}})$, $\texttt{enc}(\texttt{from}_w)$ and $\texttt{enc}(\texttt{Dom})$ (Lemma 2)
We introduce the relation $\sim$ that gathers the histories with the same predicate interpretation.

**Definition 13.** *Define the relation $\sim \subseteq H \times H$ by: for all histories $h,h' \in H$,*

$$h \sim h' \text{ iff } p^h = p^{h'}, \text{ for every } p \in P.$$

*We will denote by $[h] = \{h' \mid h \sim h'\}$ the $\sim$-equivalence class of history $h$ and by $H/\sim$ the set of all the equivalence classes, with typical element $\alpha$.*
*By definition of $\sim$, $\mathcal{I}_{h'} = \mathcal{I}_h$ for every $h' \in [h]$, which allows us to consistently define $\mathcal{I}_{[h]} := \mathcal{I}_h$ and $p^{[h]} := p^h$ for each $p \in P$.*

**Proposition 4** *$H/\sim$ is finite.*

*Proof.* We establish that there are only finitely many $p^h$ for each $p \in P$, which by definition of $\sim$ entails the finiteness of $H/\sim$. We show by induction over $h$ that each set $p^h$ belongs to a finitely generated Boolean algebra (see Appendix B for the details).

Finiteness of $H/\sim$ is a milestone in showing the automaticity of $\mathbb{H}$, as it allows us to rely on finitely many different interpretations along all the possible histories. Still, knowing that each interpretation $\mathcal{I}_h$, or equivalently $\mathcal{I}_h$ is automatic (Proposition 1) does not provide us with the mechanism to know which interpretation to consider after history $h$. The following proposition is an answer (see Appendix C for its full proof).

**Proposition 5** *For each $h \in H$, $[h]$ is regular.*

We have gathered all the material to show the automaticity of $\mathbb{H}$. We first start with the encoding of the elements of this first-order structure. Recall that because event updates are particular cases of first-order interpretations, and according to Proposition 1, each $\mathcal{I}_h$ is automatic, thus each $\mathcal{I}_{[h]}$ is automatic. Therefore there exists an automatic presentation for $\mathcal{I}_{[h]}$, over some alphabet $\Sigma_{[h]}$. We denote by $\mathtt{enc}_{[h]}$ the encoding function of this automatic presentation which maps every element of $D^{\mathcal{I}_{[h]}} = D$ onto a finite word of $\Sigma_{[h]}^*$.
Now, the overall encoding function $\mathtt{enc}$ of the domain $U = H \cup \biguplus_{h\in H} D_h$ is as follows.

**Definition 14.** *The encoding function* $\mathtt{enc} : U \to (W \cup E \cup \bigcup_{[h]\in H/\sim} \Sigma_{[h]})^*$ *is defined by:*

- *for* $h \in H$, $\mathtt{enc}(h) := h$
- *for* $d \in D_h$, $\mathtt{enc}(d) := h \cdot \mathtt{enc}_{[h]}(\mathtt{copy}_h{}^{-1}(d))$.
  *We recall that* $\mathtt{copy}_h$ *is the bijection between* $D$ *and* $D_h$.

We now prove that the encoding function $\mathtt{enc}$ of Definition 14 provides an automatic presentation of the first-order structure $\mathbb{H}$.
We recall that $\mathbb{H} = (U, (R_a^{\mathbb{H}})_{a\in\mathtt{agt}}, (p^{*\mathbb{H}})_{p\in P}, (\mathtt{from}_w^{\mathbb{H}})_{w\in W}, \mathtt{Dom}^{\mathbb{H}})$ where:

- $U = H \cup \biguplus_{h\in H} D_h$
- For $a \in \mathtt{agt}$, $R_a^{\mathbb{H}} = R_a$
- For $p \in P$, $p^{*\mathbb{H}} = \bigcup_{h\in H}(\{h\} \times p^h)$
- For $w \in W$, $\mathtt{from}_w^{\mathbb{H}} = H_w$
- $\mathtt{Dom}^{\mathbb{H}} = \bigcup_{h\in H}(\{h\} \times D_h)$

**Lemma 1.** $\mathtt{enc}(U)$ *is regular.*

*Proof.* We have that

$$\begin{aligned}
\mathtt{enc}(U) = \mathtt{enc}(H \cup \biguplus_{h\in H} D_h) &= \mathtt{enc}(H) \cup \biguplus_{h\in H} \mathtt{enc}(D_h) \\
&= H \cup \biguplus_{h\in H} h \cdot \mathtt{enc}_{[h]}(\mathtt{copy}_h{}^{-1}(D_h)) && \text{(by definition of } \mathtt{enc}) \\
&= H \cup \biguplus_{h\in H} h \cdot \mathtt{enc}_{[h]}(D) \\
&= \bigcup_{\alpha\in H/\sim} \alpha \cup \biguplus_{\alpha\in H/\sim} \alpha \cdot \mathtt{enc}_\alpha(D) && (\text{as } H = \bigcup_{\alpha\in H/\sim} \alpha)
\end{aligned}$$

In the expression above by Proposition 5, $\alpha$ is a regular language, and so is $\mathtt{enc}_\alpha(D)$ since $\mathcal{I}_\alpha$ is automatic. Moreover, by Proposition 4, the unions are finitely many, therefore $\mathtt{enc}(U)$ is a regular language.

**Lemma 2.** *Relations* $\mathtt{enc}(R_a^{\mathbb{H}})$*,* $\mathtt{enc}(p^{*\mathbb{H}})$*,* $\mathtt{enc}(\mathit{from}_w)$ *and* $\mathtt{enc}(\mathit{Dom})$ *are regular.*

*Proof.*
– $\mathtt{enc}(\mathtt{R}_a{}^{\mathbb{H}}) = \mathtt{enc}(R_a) = R_a = (H \times H) \cap R_a^0 \cdot Q_a^*$, where we recall that $R_a^0$ is the epistemic relation in $\mathcal{M}$. Now, because $W$ is finite, $R_a^0 \subseteq W \times W$ is regular. Moreover, since $E$ is finite, $Q_a \subseteq E \times E$ is regular and so is $Q_a^*$. Obviously $H \times H$ is regular. Because regular languages are closed under intersection $\mathtt{enc}(\mathtt{R}_a{}^{\mathbb{H}})$ is a regular language.
– $\mathtt{enc}(p^{*\mathbb{H}}) = \mathtt{enc}(\bigcup_{h \in H}(\{h\} \times p^h)) = \mathtt{enc}(\bigcup_{\alpha \in H/\sim} \alpha \times p^\alpha) = \bigcup_{\alpha \in H/\sim} \mathtt{enc}(\alpha) \times \mathtt{enc}(p^\alpha) = \bigcup_{\alpha \in H/\sim} \alpha \times ((\prod_{i=1}^{ar(p)} \alpha) \cdot \mathtt{enc}_\alpha(p^\alpha))$. Because regular languages are closed under cartesian product and union, this last expression describes a regular language.
– $\mathtt{enc}(\mathtt{from}_w) = \mathtt{enc}(H_w) = H_w = H \cap w \cdot E^*$. which is clearly a regular language.
– $\mathtt{enc}(\mathtt{Dom}\ ^{\mathbb{H}}) = \mathtt{enc}(\bigcup_{h \in H}(\{h\} \times D_h)) = \bigcup_{\alpha \in H/\sim} \alpha \times \mathtt{enc}_\alpha(D)$. As a finite union of regular languages, this last expression describes a regular language.

Lemmas 1 and 2 conclude the proof of Theorem 5.

## 4 Conclusion and Future work

Inheriting from DEL, the general Epistemic Planning Problem in the DFOEL setting (EPP) with possibly modal pre-conditions is undecidable. Furthermore, we discovered that first-order quantifications in (non-modal) post-conditions draws a line between decidability and undecidability. Our proof for the decidable case (EPP0) involves a non-trivial machinery based on automatic structures. Even though one might wish a simpler proof, the chosen approach offers a wide range of decidable problems in the EPP0 setting. Indeed, in the same line as [5], the proof of automaticity allows one to finitely represent the set of all plan solutions and to address all sorts of queries on it such as their infinity. On a longer term perspective, we may consider, as done in [13], event schemes in action models resulting in an infinity of events for infinite domains; it is reasonable to conjecture that our results on decidability would still hold true in this case. Another possible track would be to allow each world to have a different domains.

## References


1. Guillaume Aucher, Bastien Maubert, and Sophie Pinchinat. Automata techniques for epistemic protocol synthesis. In Fabio Mogavero, Aniello Murano, and Moshe Y. Vardi, editors, *Proceedings 2nd International Workshop on Strategic Reasoning, SR 2014, Grenoble, France, April 5-6, 2014*, volume 146 of *EPTCS*, pages 97–103, 2014.
2. Patrick Blackburn, Johan Van Benthem, and Frank Wolter. *Handbook of Modal Logic*. Springer, 2006.

3. Achim Blumensath and Erich Grädel. Automatic structures. In *Proceedings of the 15th Annual IEEE Symposium on Logic in Computer Science*, LICS '00, page 51, USA, 2000. IEEE Computer Society.
4. Thomas Bolander and Mikkel Birkegaard Andersen. Epistemic planning for single and multi-agent systems. *J. Appl. Non Class. Logics*, 21(1):9–34, 2011.
5. Thomas Bolander, Tristan Charrier, Sophie Pinchinat, and François Schwarzentruber. Del-based epistemic planning: Decidability and complexity. *Artificial Intelligence*, 287:103304, 2020.
6. Tristan Charrier, Bastien Maubert, and François Schwarzentruber. On the Impact of Modal Depth in Epistemic Planning (Extended Version). Research report, IRISA, équipe LogicA, May 2016.
7. Sébastien Lê Cong, Sophie Pinchinat, and François Schwarzentruber. Small undecidable problems in epistemic planning. In Jérôme Lang, editor, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden*, pages 4780–4786. ijcai.org, 2018.
8. Wilfrid Hodges et al. *A shorter model theory.* Cambridge university press, 1997.
9. John E Hopcroft, Rajeev Motwani, and Jeffrey D Ullman. Introduction to automata theory, languages, and computation. *Acm Sigact News*, 32(1):60–65, 2001.
10. Bakhadyr Khoussainov, André Nies, Sasha Rubin, and Frank Stephan. Automatic structures: Richness and limitations. *Log. Methods Comput. Sci.*, 3(2), 2007.
11. Barteld Kooi. Dynamic term-modal logic. *Van Benthem, J., Ju, S. and Veltman, F. (eds.). A meeting of the minds. Proceedings of the workshop on Logic, Rationality and Interaction, Beijing , 2007 Texts in Computing Computer Science 8. pp. 173-185 College publications , London*, 8:173–185, 2008.
12. Bastien Maubert. *Logical foundations of games with imperfect information : uniform strategies.* Theses, Université Rennes 1, January 2014.
13. Andrés Occhipinti Liberman, Andreas Achen, and Rasmus Kræmmer Rendsvig. Dynamic term-modal logics for first-order epistemic planning. *Artificial Intelligence*, 286:103305, 2020.
14. Andrés Occhipinti Liberman and Rasmus Kræmmer Rendsvig. Decidability results in first-order epistemic planning. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 4161–4167. International Joint Conferences on Artificial Intelligence Organization, 7 2020. Main track.
15. Sasha Rubin. Automata presenting structures: A survey of the finite string case. *The Bulletin of Symbolic Logic*, 14(2):169–209, 2008.
16. Hans Van Ditmarsch, Wiebe van Der Hoek, and Barteld Kooi. *Dynamic epistemic logic*, volume 337. Springer Science & Business Media, 2007.

## Appendix A Proof of Theorem 3

Let $M = (Q, \Gamma, q_0, \delta, F)$ be a Turing machine. We consider the first-order structure over the signature with a binary predicate $p$ and two unary predicates $i$ and $f$ defined by: $\mathcal{G} := (C(M), p^{\mathcal{G}}, i^{\mathcal{G}}, f^{\mathcal{G}})$ where $C(M)$ is the set of configurations of $M$, $p^{\mathcal{G}}$ is the relation $\to$ composed of configuration pairs such that $M$ can move from the former to the latter, $i^{\mathcal{G}}$ is the set of initial configurations, and $f^{\mathcal{G}}$ is the set of final configurations.
As any first-order structure, the structure $\mathcal{G}$ can be embarked into a single-agent single-world epistemic model $\mathcal{M}_{\mathcal{G}} = (\{w\}, \{(w, w)\}, \mathcal{G})$.

**Lemma 3.** *The epistemic model $\mathcal{M}_{\mathcal{G}}$ is automatic.*

*Proof.* We already know that the configuration graph of a Turing machine is automatically presentable [10, Lemma 5.1]. Moreover, with the encoding of the structure $(C(M), \to)$ used to prove this fact, one can easily verify that two unary relations $i^w$ and $f^w$ are also regular, which concludes.

We next tune a single-event action model $\mathcal{E} = (\{e\}, \{(e, e)\}, pre, post)$ whose effect is to enrich the current binary relation between configurations with its self-composition.

- $pre(e) = \top$
- $post(e)(p)(z_1, z_2) = p(z_1, z_2) \vee \exists x(p(z_1, x) \wedge p(x, z_2))$, and
- $post(e)(i)(z_1) = i(z_1)$ and $post(e)(f)(z_1) = f(z_1)$, i.e., the initial and final configurations remain unchanged.

Finally, we define the goal formula $\exists x \exists y(i(x) \wedge p(x, y) \wedge f(y))$ which states the existence of an initial and a final configuration related by $p$.
We claim that we indeed have defined a reduction from TME to EPP1. Indeed, by definition of $post(e)(p)$, a sequence $ee \dots e$ of updates makes predicate $p$ incrementally get closer to the transitive closure of $\to$. More precisely, one can easily show by induction over $\ell$ that after $\ell$ triggers of event $e$, two configurations are related by the interpretation of $p$ if and only if there exists a path of length at most $\ell$ between them. Therefore, the goal formula $\exists x \exists y(i(x) \wedge p(x, y) \wedge f(y))$ eventually holds after finitely many triggers of event $e$ if and only if some final configuration is reachable from some initial configuration in the initial model $\mathcal{G}$, otherwise said, if and only if the language of the Turing machine is non-empty.

## Appendix B Proof of Proposition 4

We will prove that every possible interpretation $p^h$ is obtained by boolean combination of a finite set of atoms obtained from the interpretations $q^w$. We first describe how to obtain those atoms with Definition 15, then we show that the interpretation of a predicate after an update is a combination of atoms from the precedent interpretations (Lemma 4). Finally, we use Lemma 5 for an induction on the histories in Lemma 6. For $n \in \mathbb{N}$, denote by $[\![n]\!]$ the set $\{1, ..., n\}$.

**Definition 15.** *Let $k, \ell \in \mathbb{N}$ and $A \subseteq D^k$, if $\sigma : [\![k]\!] \to [\![\ell]\!]$ define $A\sigma = \{(d_1, ..., d_\ell) \mid (d_{\sigma(1)}, ..., d_{\sigma(k)}) \in A\}$*

For a set of generators $G$ denote by $\mathbb{B}G$ the boolean algebra generated by $G$.

**Lemma 4.** *For $he \in H$, and $p \in P$, $p^{he} \in \mathbb{B}\{q^h\sigma \mid q \in P,\ \sigma : [\![ar(q)]\!] \to [\![ar(p)]\!]\}$*

*Proof.* If $post(e)(p) = q(z_{\sigma(1)}, ..., z_{\sigma(ar(q))})$ we have that:
$(d_1, ..., d_{ar(p)}) \in p^{he}$ iff $\mathcal{I}_h \models_{[z_1 \mapsto d_1, ..., z_{ar(p)} \mapsto d_{ar(p)}]} q(z_{\sigma(1)}, ..., z_{\sigma(ar(q))})$
iff $(d_{\sigma(1)}, ..., d_{\sigma(ar(q))}) \in q^h$
iff $(d_1, ..., d_{ar(p)}) \in q^h\sigma$
Therefore $p^{he} = q^h\sigma$.
Thus if $post(e)(p) \in \mathcal{L}_0$, then $p^{he} \in \mathbb{B}\{q^h\sigma \mid q \in P,\ \sigma : [\![ar(q)]\!] \to [\![ar(p)]\!]\}$.

**Lemma 5.** *Let $k, \ell, n \in \mathbb{N}$, if $A \subseteq D^k$, let $\sigma : [\![\ell]\!] \to [\![n]\!]$ and $\nu : [\![k]\!] \to [\![\ell]\!]$,*
*then $(A\nu)\sigma = A(\sigma \circ \nu)$*

*Proof.* We have that:
$(d_1, ..., d_{ar(p)}) \in (A\nu)\sigma$ iff $(d_{\sigma(1)}, ..., d_{\sigma(\ell)}) \in A\nu$
iff $(d_{\sigma(\nu(1))}, ..., d_{\sigma(\nu(k))}) \in A$
iff $(d_1, ..., d_{ar(p)}) \in A(\sigma \circ \nu)$

**Lemma 6.** *For $p \in P$, and $h \in H$, $p^h \in \mathbb{B}\{q^w\sigma \mid q \in P,\ \sigma : [\![ar(q)]\!] \to [\![ar(p)]\!]\}$*

*Proof.* By induction on $h$, define the induction hypothesis: $\mathcal{P}(h)$ : "$\forall p \in P$, $p^h \in \mathbb{B}\{q^w\sigma \mid q \in P,\ \sigma : [\![ar(q)]\!] \to [\![ar(p)]\!]\}$".
For $w \in W$, and $p \in P$, $p^w = p^w id$.
Let $he \in H$, if $\mathcal{P}(h)$, then by lemma 4, we have that $p^{he} \in \mathbb{B}\{q^h\sigma \mid q \in P,\ \sigma : [\![ar(q)]\!] \to [\![ar(p)]\!]\}$.
If $q \in P$, and $\sigma : [\![ar(q)]\!] \to [\![ar(p)]\!]$, we have by $\mathcal{P}(h)$, $q^h \in \mathbb{B}\{r^w\nu \mid r \in P,\ \nu : [\![ar(r)]\!] \to [\![ar(q)]\!]\}$
Thus $q^h\sigma \in \mathbb{B}\{(r^w\nu)\sigma \mid r \in P,\ \nu : [\![ar(r)]\!] \to [\![ar(q)]\!]\}$
Then by lemma 5, we have that $q^h\sigma \in \mathbb{B}\{r^w(\sigma \circ \nu) \mid r \in P,\ \nu : [\![ar(r)]\!] \to [\![ar(q)]\!]\}$
i.e., $q^h\sigma \in \mathbb{B}\{r^w\nu \mid r \in P,\ \nu : [\![ar(r)]\!] \to [\![ar(p)]\!]\}$
Therefore $p^{he} \in \mathbb{B}\{q^w\sigma \mid q \in P,\ \sigma : [\![ar(r)]\!] \to [\![ar(p)]\!]\}$

$\{q^w\sigma \mid q \in P,\ \sigma : [\![ar(r)]\!] \to [\![ar(p)]\!]\}$ is finite, so that $\mathbb{B}\{q^w\sigma \mid q \in P,\ \sigma : [\![ar(r)]\!] \to [\![ar(p)]\!]\}$ is also finite. Therefore there is a finite number of different $p^h$.

## Appendix C Proof of Proposition 5

Before proving Proposition 5, we introduce some material and prove Lemma 7.

**Definition 16.** *Let $\mathcal{I}$ be a $P$-interpretation with domain $D$, and $e \in E$, if $\mathcal{I} \models pre(e)$, we define the $P$-interpretation:*

$$\mathcal{I} \otimes e := (D, (p^{\mathcal{I} \otimes e})_{p \in P})$$

*with* $p^{\mathcal{I} \otimes e} := \{(d_1, ..., d_{kp}) \mid \mathcal{I} \models_{[z_1 \mapsto d_1, ..., z_{ar(p)} \mapsto d_{ar(p)}]} post(e)(p)(z_1, ..., z_{ar(p)})\}$, *for each* $p \in P$.

**Lemma 7.** *For* $he \in H$, $\mathcal{I}_{[he]} = \mathcal{I}_{[h]} \otimes e$

*Proof.* As $post(e)(p)$ is non-modal, we have that

$$H, h \models_{[z_1 \mapsto d_1, ..., z_{ar(p)} \mapsto d_{ar(p)}]} post(e)(p)(z_1, ..., z_{kp})$$
if, and only if,
$$\mathcal{I}_h \models_{[z_1 \mapsto d_1, ..., z_{ar(p)} \mapsto d_{ar(p)}]} post(e)(p)(z_1, ..., z_{kp}).$$

We now turn to proving Proposition 5 about the regularity of $[h]$ by constructing a finite-state automaton for it: Let $\mathcal{A}_{[h]}$ be the automaton over alphabet $W \cup E$ with states ranging over $\{s_0\} \cup H/\sim$, initial and final states ranging over $\{s_0\}$ and $\{[h]\}$ respectively, and whose transition relation $\delta$ is defined by:

- For $w \in W$, $\delta(s_0, w) = [w]$;
- For $\alpha \in H/\sim$ and $e \in E$, if $\mathcal{I}_\alpha \models pre(e)$, then $\delta(\alpha, e) = \alpha'$ s.t. $\mathcal{I}_{\alpha'} = \mathcal{I}_\alpha \otimes e$.

We argue that $L(\mathcal{A}_{[h]}) = [h]$:

$$\begin{aligned} we_1...e_n \in [h] &\text{ iff } \mathcal{I}_{[we_1...e_n]} = \mathcal{I}_{[h]} \\ &\text{ iff } \mathcal{I}_{[we_1...e_n]} \text{ is final} \\ &\text{ iff } \mathcal{I}_{[w]} \otimes e_1 \otimes ... \otimes e_n \text{ is final.} \\ &\text{ iff } \delta^*(s_0, we_1...e_n) \text{ is final} \\ &\text{ iff } we_1...e_n \in L(\mathcal{A}_{[h]}) \end{aligned}$$

This achieves the proof of Proposition 5.