

Reachability Games in Dynamic Epistemic Logic

Bastien Maubert¹, Sophie Pinchinat² and François Schwarzentruber²

¹Università degli Studi di Napoli “Federico II”, Italy

²Univ Rennes, CNRS, IRISA, France

Abstract

We define reachability games based on Dynamic Epistemic Logic (DEL), where the players’ actions are finely described as DEL action models. We first consider the setting where an external controller with perfect information interacts with an environment and aims at reaching some desired state of knowledge regarding the passive agents of the system. We study the problem of strategy existence for the controller, which generalises the classic epistemic planning problem, and we solve it for several types of actions such as public announcements and public actions. We then consider a yet richer setting where agents themselves are players, whose strategies must be based on their observations. We establish several (un)decidability results for the problem of existence of a distributed strategy, depending on the type of actions the players can use, and relate them to results from the literature on multiplayer games with imperfect information.

1 Introduction

Many applications fall within the scope of reachability games with imperfect information, such as video games [Coulombe and Lynch, 2018], *Kriegspiel* (the epistemic variant of Chess) [Matros, 2018], Hanabi [Baffier *et al.*, 2016], or contingent and conformant planning [Geffner and Bonet, 2013].

Games with imperfect information are computationally hard, and even undecidable for multiple players [Peterson and Reif, 1979]. One way to tame this complexity is to make assumptions on how the knowledge of the different players compare: if all players that cooperate can be ordered in a hierarchy where one knows more than the next, a situation called *hierarchical information*, then the existence of distributed strategies can be decided [Peterson *et al.*, 2002; Berwanger *et al.*, 2018]. Another natural approach is to consider fragments based on classes of action types, as done for instance in [Ramanujam and Simon, 2010; Belardinelli *et al.*, 2017b; Bouyer, 2018] where different kinds of public actions are considered. But the usual graph-based models of games with imperfect information, where the players’ actions are modelled as labels on the edges, make it difficult to define subtle properties of actions.

By contrast, Dynamic epistemic logic (DEL) [van Ditmarsch *et al.*, 2008] was designed to describe actions precisely: how they affect the world and how they are perceived. In particular, classic action types such as public/private announcements or public actions correspond to natural classes of DEL *action models*. Also, DEL extends epistemic logic and hence enables modelling *higher-order knowledge*, i.e. what an agent knows about what another agent knows etc, and the evolution of agents’ knowledge over time.

A classification of the complexity with respect to action types was addressed in the literature of *epistemic planning*, a problem that asks for the existence of a *plan*, i.e. a finite sequence of DEL actions to reach a situation that satisfies some given objective expressed in epistemic logic. However this problem, which can be seen as solving one-player reachability games with epistemic objective, has never been considered in a strategic, adversarial context. Our work bridges the gap between DEL and games by introducing adversarial aspects in DEL planning, thus moving from plan generation to strategy synthesis. We define two frameworks for DEL-based reachability games, where players start in a given epistemic situation and their possible moves are described by action models, and the objective is given as an epistemic formula.

In a first step we consider *open systems* [Harel and Pnueli, 1985], i.e. systems that interact with an environment. In our setting, two omniscient, external entities (the *controller* and the *environment*) choose in turn which actions are performed. We call this setting *DEL controller synthesis*. Here, agents involved in the models and formulas are not active, they merely observe how the system evolves based on the actions chosen by the controller and the environment, and update their knowledge accordingly. DEL controller synthesis extends DEL planning, as the latter is a degenerate case of the former where the environment stays idle, and we therefore inherit undecidability for the general case. Nevertheless we show that, as for DEL planning, decidability is regained when actions do not increase uncertainty (so-called *non-expanding actions*) or when the preconditions of actions are propositional formulas. More precisely, we show PSPACE-completeness when possible moves are public announcements, EXPTIME-completeness for the more general public actions, and membership in $(k + 1)$ -EXPTIME for propositional actions when the objectives are formulas of modal depth at most k .

We then generalise further this setting by turning agents

	Public announcements	Public actions	Propositional actions	Full
Plan	NP-complete	PSPACE-complete	decidable	undecidable
Controller	PSPACE-complete (Th. 2)	EXPTIME-complete (Th. 3)	decidable (Th. 4)	undecidable
Distributed strategy	PSPACE-complete (Th. 8)	EXPTIME-complete (Th. 9)	undecidable (Th. 7) decidable case (Th. 10)	undecidable

Table 1: Known and new results (in grey) of the plan, controller and distributed strategy synthesis problems.

into players. Unlike the omniscient controller of the former setting, agents have imperfect information about the current state of the game, and can only base their decisions on what they know. In the theory of games with imperfect information this is modelled by the notion of *uniform strategies*, also called *observation-based strategies* [Apt and Grädel, 2011]. We study the problem of *distributed strategy synthesis*, where a group of players cooperate to enforce some objective against the remaining players. As for multiplayer games with imperfect information the problem is undecidable, already for propositional actions and a coalition of two players. However we show that the two kinds of assumptions that make imperfect-information games decidable, namely public actions and hierarchical information, also yield decidable cases of multiplayer DEL games. Furthermore, in the case of public announcements and public actions, the complexity is not worse than for controller synthesis.

Table 1 sums up previous results for epistemic planning, as well as the results established in this contribution. More detailed proofs can be found in [Maubert *et al.*, 2019].

2 Related Work

The complexity of DEL-based epistemic planning has been thoroughly investigated. It is undecidable already for actions with preconditions of modal depth one and propositional postconditions [Bolander and Andersen, 2011; Lê Cong *et al.*, 2018]. For preconditions of modal depth one and no postconditions the problem has been open for years, but it is decidable when pre- and postconditions are propositional [Yu *et al.*, 2013; Aucher *et al.*, 2014; Douéneau-Tabot *et al.*, 2018]. It is also known to be NP-complete for public announcements [Bolander *et al.*, 2015; Charrier, 2018], and PSPACE-complete for public actions [Charrier, 2018].

The decidability for propositional actions has been extended in [Aucher *et al.*, 2014] by considering infinite trees of actions called *protocols* instead of finite plans, and specifications in branching-time epistemic temporal logic instead of reachability for epistemic formulas; this has been extended further in [Douéneau-Tabot *et al.*, 2018] by enriching the specification language with Chain Monadic Second-order Logic. Both results rely on the fact that when actions are propositional, the infinite structures generated by repeated application of action models form a class of *regular structures* [Aucher *et al.*, 2014; Maubert, 2014], i.e. relational structures that have a finite representation via automata. First-order logic is decidable on such structures [Blumensath and Grädel, 2000], and chain-MSO is decidable on a subclass called *regular automatic trees* [Douéneau-Tabot *et al.*,

2018], but neither of these logics can express the existence of strategies in games. However we will show that the regular structures obtained from propositional DEL models can be seen as finite turn-based game arenas studied in games played on graphs. This allows us to transfer decidability results on games with epistemic temporal objectives to the DEL setting.

A notion of cooperative planning in DEL has been studied in [Engesser *et al.*, 2017], but without the adversarial aspect of games. Also, in [Lima, 2014], a game setting has been developed with the so-called Alternating-time Temporal Dynamic Epistemic Logic, but it does not consider uniform strategies and thus cannot express the existence of distributed strategies. Our controller synthesis problem can be expressed in this logic, but not in the fragment that they solve, which cannot express reachability. On the other hand, several decidability results for logics for strategic and epistemic reasoning have been established recently [Belardinelli *et al.*, 2017a; Maubert and Murano, 2018], but they do not offer the fine modelling of actions possible in DEL. For instance they cannot easily model public announcements, which we show yield better complexity than those obtained in their settings.

3 Background in Epistemic Planning

Let us fix a countable set of *atomic propositions* AP .

3.1 The Classic DEL Setting

We recall models of epistemic logic [Fagin *et al.*, 1995].

Definition 1 An epistemic model $\mathcal{M} = (W, (R_a)_{a \in \text{Agt}}, V)$ is a tuple where W is a non-empty finite set of possible worlds (or situations), $R_a \subseteq W \times W$ is an accessibility relation for agent a , $V : W \rightarrow 2^{AP}$ is a valuation function.

We write $wR_a u$ instead of $(w, u) \in R_a$; its intended meaning is that when the actual world is w , agent a considers that u may be the actual world. The valuation function V provides the subset of atomic propositions that hold in a world. A pair (\mathcal{M}, w) is called a *pointed epistemic model*, and we let $|\mathcal{M}|$ be the *size* of \mathcal{M} , defined as $|W| + \sum_{a \in \text{Agt}} |R_a| + \sum_{w \in W} |V(w)|$. We will only consider finite models, i.e. we assume that $V(w)$ is finite for all worlds.

The syntax of Epistemic Logic \mathcal{L}_{EL} is given by the following grammar: $\varphi ::= p \mid \neg\varphi \mid (\varphi \vee \varphi) \mid K_a\varphi$, where p ranges over AP and a ranges over Agt .

$K_a\varphi$ is read ‘agent a knows that φ is true’. We define the usual abbreviations $(\varphi_1 \wedge \varphi_2)$ for $\neg(\neg\varphi_1 \vee \neg\varphi_2)$ and $\hat{K}_a\varphi$ for $\neg K_a\neg\varphi$, and use $\mathcal{L}_{\text{Prop}}$ for the fragment of \mathcal{L}_{EL} with propositional formulas only. The *modal depth* of a formula is its maximal number of nested knowledge operators; for instance,

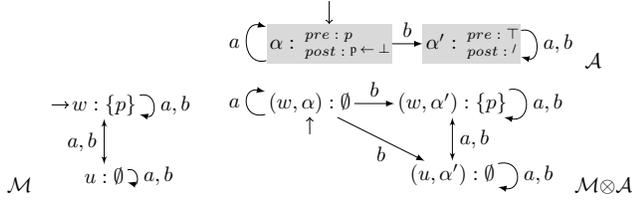


Figure 1: Example of product. Symbol \perp indicates the trivial postcondition that leaves valuations unchanged.

the formula $K_a K_b p \wedge \neg K_a q$ has modal depth 2. The *size* $|\varphi|$ of a formula φ is the number of symbols in it.

The semantics of \mathcal{L}_{EL} relies on *pointed epistemic models*. We define $\mathcal{M}, w \models \varphi$, read as ‘formula φ holds in the pointed epistemic model (\mathcal{M}, w) ’, by induction on φ with expected meaning for propositional operators and $\mathcal{M}, w \models K_a \varphi$ if for all u such that $w R_a u$, $\mathcal{M}, u \models \varphi$.

Dynamic Epistemic Logic (DEL) relies on *action models* (also called ‘‘event models’’), that specify how agents perceive the occurrence of an action as well as its effects on the world.

Definition 2 An action model $\mathcal{A} = (A, (R_a^A)_{a \in \text{Agt}}, \text{pre}, \text{post})$ is a tuple where: A is a non-empty finite set of possible actions, $R_a^A \subseteq A \times A$ is the accessibility relation for agent a , $\text{pre} : A \rightarrow \mathcal{L}_{\text{EL}}$ provides the precondition for an action to be performed, and $\text{post} : A \times AP \rightarrow \mathcal{L}_{\text{Prop}}$ provides the postcondition (i.e. the effects) of an action.

A *pointed action model* is a pair (\mathcal{A}, α) where α represents the actual action. We let $|\mathcal{A}|$ be the *size* of \mathcal{A} , defined as $|\mathcal{A}| := |A| + \sum_{a \in \text{Agt}} |R_a^A| + \sum_{\alpha \in A} |\text{pre}(\alpha)| + \sum_{\alpha \in A, p \in AP} |\text{post}(\alpha, p)|$.

An action α is *executable* in a world w of an epistemic model \mathcal{M} if $\mathcal{M}, w \models \text{pre}(\alpha)$, and in that case we define $V(w, \alpha) := \{p \in AP \mid \mathcal{M}, w \models \text{post}(\alpha, p)\}$, the set of atomic propositions that hold after occurrence of action α in world w . Since postconditions are always propositional, we can define similarly $V(v, \alpha)$ where $v \subseteq 2^{AP}$ is a valuation.

Types of actions. We identify noticeable types of actions. An action model \mathcal{A} is *propositional* if all pre- and postconditions of actions in \mathcal{A} belong to $\mathcal{L}_{\text{Prop}}$. A *public action* is a pointed action model \mathcal{A}, α such that for each agent a , R_a^A is the identity relation. A *public announcement* is a public action \mathcal{A}, α such that for all p , $\text{post}(\alpha, p) = p$.

We recall the product that models how to update an epistemic model when an action is executed [Baltag *et al.*, 1998].

Definition 3 Let $\mathcal{M} = (W, (R_a)_{a \in \text{Agt}}, V)$ be an epistemic model, and $\mathcal{A} = (A, (R_a^A)_{a \in \text{Agt}}, \text{pre}, \text{post})$ be an action model. The product of \mathcal{M} and \mathcal{A} is defined as $\mathcal{M} \otimes \mathcal{A} = (W', (R_a)_{a \in \text{Agt}}, V')$ where:

- $W' = \{(w, \alpha) \in W \times A \mid \mathcal{M}, w \models \text{pre}(\alpha)\}$;
- $(w, \alpha) R'_a (w', \alpha')$ if $w R_a w'$ and $\alpha R_a^A \alpha'$;
- $V'((w, \alpha)) = V(w, \alpha)$.

Example 1 Figure 1 shows the pointed model \mathcal{M}, w that represents a situation in which p is true and both agents a and b

do not know it. The pointed action model \mathcal{A}, α describes the action where agent a learns that p was true but that it is now set to false, while agent b does not learn anything (she sees action α' that has trivial pre- and postcondition). In the product epistemic model $\mathcal{M} \otimes \mathcal{A}, (w, \alpha)$, agent a now knows that p is false, while b still does not know the truth value of p , or whether agent a knows it.

An epistemic model (resp. an action model) is *S5* if all accessibility relations are equivalence relations. This property is important to model games with imperfect information, and we will assume it in Section 5.

3.2 Generated Structure

Iteratively executing an action model from an initial epistemic model generates an infinite sequence of epistemic models, whose union yields an infinite epistemic structure where dynamics are represented by the possible sequences of actions, while information is captured by the accessibility relations.

Definition 4 Given $\mathcal{M} = (W, \{R_a\}_{a \in \text{Agt}}, V)$ an epistemic model and $\mathcal{A} = (A, \{R_a^A\}_{a \in \text{Agt}}, \text{pre}, \text{post})$ an action model, we define the family of disjoint epistemic models $\{\mathcal{M}\mathcal{A}^n\}_{n \geq 0}$ by letting $\mathcal{M}\mathcal{A}^0 = \mathcal{M}$ and $\mathcal{M}\mathcal{A}^{n+1} = \mathcal{M}\mathcal{A}^n \otimes \mathcal{A}$. We finally define the infinite epistemic model $\mathcal{M}\mathcal{A}^* = \bigcup_{n \in \mathbb{N}} \mathcal{M}\mathcal{A}^n$.

Anticipating the game setting we later define, we call a *play* an infinite sequence $\pi = w\alpha_1\alpha_2\dots$ such that all finite prefixes of π are in $\mathcal{M}\mathcal{A}^*$. A *history* is a finite prefix h of a play. We let $\text{Plays}_{\mathcal{M}\mathcal{A}^*}(w)$ and $\text{Hist}_{\mathcal{M}\mathcal{A}^*}(w)$ be, respectively, the set of all plays and histories in $\mathcal{M}\mathcal{A}^*$ that start with w . These definitions entail the following.

Lemma 1 For every world $(w, \alpha_1, \dots, \alpha_n) \in \mathcal{M} \otimes \mathcal{A}^n$, and every formula $\varphi \in \mathcal{L}_{\text{EL}}$,

$$\mathcal{M} \otimes \mathcal{A}^n, (w, \alpha_1, \dots, \alpha_n) \models \varphi \text{ iff } \mathcal{M}\mathcal{A}^*, w\alpha_1\dots\alpha_n \models \varphi.$$

This shows that the alternative definition of epistemic planning given in the next section is equivalent to the usual one.

3.3 Epistemic Planning

The epistemic planning problem asks for the existence of an executable sequence of designated actions $\alpha_1, \dots, \alpha_n$ in an action model \mathcal{A} , whose execution from \mathcal{M}, w leads to a situation satisfying some objective expressed as an epistemic logic formula¹. Formally, we consider the following problem.

Definition 5 (Plan existence problem)

- *Input:* a pointed epistemic model \mathcal{M}, w , an action model \mathcal{A} and an objective formula $\varphi \in \mathcal{L}_{\text{EL}}$;
- *Output:* yes if there is a history h in $\text{Hist}_{\mathcal{M}\mathcal{A}^*}(w)$ such that $\mathcal{M}\mathcal{A}^*, h \models \varphi$.

Main known results on the plan existence problem are summarised in Table 1, while the relevant pointers to the literature are given in the related work paragraph of the introduction.

In the rest of this paper, we will need standard notions and notations that we recall here. A *finite* (resp. *infinite*) word

¹This formulation of the problem is equivalent to the usual one with n action models: they are interreducible in linear time.

over some alphabet Σ is an element of Σ^* (resp. Σ^ω). The length of a finite word $w = w_0w_1 \dots w_n$ is $|w| := n + 1$, and $\text{last}(w) := w_n$ is its last letter. Given a finite (resp. infinite) word w and $0 \leq i < |w|$ (resp. $i \in \mathbb{N}$), we let w_i be the letter at position i in w , $w_{\leq i} := w_0 \dots w_i$ is the prefix of w that ends at position i and $w_{\geq i} := w_iw_{i+1} \dots$ is the suffix of w that starts at position i . We also use variables x that range over some finite domain. We will write $(x = d)$ for the fact “the value of x is d ”, and use $x := d$ for the effect of setting x to value d . This can all be encoded with atomic propositions.

4 Controller Synthesis

We first generalise the plan existence problem to the setting where some environment may perturb the execution of the plan that should thus be robust against it.

Formally, we consider an initial epistemic model \mathcal{M} , as in Definition 1, with an initial world w_i , and an action model $\mathcal{A} = (A, (R_a^A)_{a \in \text{Agt}}, \text{pre}, \text{post})$ whose set of actions A is partitioned into actions in A_{ctr} controlled by a *Controller* and actions in A_{env} controlled by the *Environment*.

Controller and Environment play in turn: in each round, Controller first chooses to execute an action in A_{ctr} , then Environment chooses to execute an action in A_{env} . Thus instead of seeking a history in $\mathcal{M}\mathcal{A}^*$ that reaches an objective formula, as in epistemic planning, one seeks a *strategy* for Controller: formally, it is a partial function $\sigma : \text{Hist}_{\mathcal{M}\mathcal{A}^*}(w_i) \rightarrow \text{A}_{\text{ctr}}$ defined on histories of odd length (when it is the controller’s turn). An *outcome* of a strategy σ is a play $\pi = w_i\alpha_1\alpha_2 \dots$ in which the controller follows σ , i.e. for all $i \in \mathbb{N}$, $\alpha_{2i+1} = \sigma(\pi_{\leq 2i}) \in \text{A}_{\text{ctr}}$, while the other actions, of the form α_{2i+2} , are selected by the environment. A strategy σ for Controller is *winning* for an objective formula $\varphi \in \mathcal{L}_{\text{EL}}$ if for every outcome π of σ , there exists $i \in \mathbb{N}$ s.t. $\mathcal{M}\mathcal{A}^*, \pi_{\leq i} \models \varphi$.

Definition 6 (The controller synthesis problem)

- *Input:* a pointed epistemic model \mathcal{M}, w_i , action model \mathcal{A} with $A = \text{A}_{\text{ctr}} \uplus \text{A}_{\text{env}}$, and an objective $\varphi \in \mathcal{L}_{\text{EL}}$;
- *Output:* yes if there exists a winning strategy for Controller for objective φ ; no otherwise.

Remark 1 Formally, we define and study the problem of existence of a strategy. We take the liberty to call the problem controller synthesis because all the algorithms we provide can produce a winning strategy whenever there exists one. The same remark applies to the distributed strategy synthesis problem defined in the next section.

As the plan existence problem reduces to the controller synthesis problem, the undecidability of the former entails the one of the latter. We next establish that in all known subcases where the plan existence problem is decidable, so is the controller synthesis problem.

4.1 The Case of Non-Expanding Action Models

We consider so-called *non-expanding* action models where actions do not expand epistemic models when executed, like public actions. For this type of actions, the search space is finite and thus the problem is decidable. We establish the precise computational complexity of the problem in these cases.

Theorem 2 *When actions are public announcements, the controller synthesis problem is PSPACE-complete.*

Proof sketch. Since applying public announcements to epistemic models only removes worlds, and does not change those that remain, the number of successive public announcements to consider can be bounded by the number of worlds in the initial epistemic model. We can thus solve the problem with an alternating algorithm that runs in polynomial time, guessing existentially actions of the controller and universally those of the environment. We conclude by recalling that alternating polynomial time corresponds to deterministic polynomial space [Chandra and Stockmeyer, 1976]. PSPACE-hardness is proven by reduction from QBF. \square

Theorem 3 *When actions are public, the controller synthesis problem is EXPTIME-complete.*

Proof sketch. As for public announcements, applying a public action does not add worlds but may change facts, so that linear-size sequences of actions may not suffice. Nonetheless, we can turn the alternating algorithm from the proof of Theorem 2 into one that runs in polynomial-space. The EXPTIME-membership follows from the fact that alternating polynomial space corresponds to exponential time [Chandra and Stockmeyer, 1976]. EXPTIME-hardness is obtained by reduction from the conditional planning problem, a variant of classical planning with non-deterministic actions [Littman *et al.*, 1998; Rintanen, 2004]. \square

Theorem 3 also generalises to other non-expanding actions models such as the so-called *separable action models* [Charrier, 2018], where the preconditions of any two actions in the same connected component are logically inconsistent.

4.2 The Case of Propositional Action Models

To solve our controller synthesis problem we rely on the approach followed in [Maubert, 2014] to solve the plan existence problem for propositional actions. This approach has two main ingredients: (I1) when \mathcal{A} is propositional, the generated structure $\mathcal{M}\mathcal{A}^*$ can be represented finitely, and (I2) one can decide the existence of a winning strategy in a certain class of two-player games with epistemic objectives.

Theorem 4 *When action models are propositional, the controller synthesis problem is decidable, and in $(k + 1)$ -EXPTIME if the objective’s modal depth is bounded by k .*

We devote the rest of this section to prove Theorem 4, which requires to introduce particular game arenas.

Definition 7 *A two-player epistemic game arena is a structure $\mathcal{G} = (W, w_i, \Delta, (R_a)_{a \in \text{Agt}}, V)$ where $(W, (R_a)_{a \in \text{Agt}}, V)$ is an epistemic model, $W = W_0 \uplus W_1$ is partitioned into the positions of players 0 and 1, w_i is an initial world and $\Delta \subseteq W \times W$ is a transition relation.*

A *play* in a game arena \mathcal{G} is an infinite sequence of worlds $\pi = w_0w_1w_2 \dots$ such that for all $i \in \mathbb{N}$, $w_i \Delta w_{i+1}$, and a *history* is a finite nonempty prefix of a play. We let $\text{Plays}_{\mathcal{G}}$ and $\text{Hist}_{\mathcal{G}}$ be the sets of plays and histories in \mathcal{G} , respectively. Accessibility relations $(R_a)_{a \in \text{Agt}}$ are extended to histories to interpret epistemic formulas: two histories $h = w_0 \dots w_n$

and $h' = w'_0 \dots w'_m$ are related by R_a whenever $n = m$ (same length) and $w_i R_a w'_i$ for every $i \leq n$.

A strategy for player 0 is a partial function $\sigma : \text{Hist}_{\mathcal{G}} \rightarrow W$ such that for every h with $\text{last}(h) \in W_0$: $\text{last}(h) \Delta \sigma(h)$. A play $\pi = w_1 w_2 \dots$ is an *outcome* of σ if for every $i \in \mathbb{N}$ with $\pi_i \in W_1$, we have $\pi_{i+1} = \sigma(\pi_{\leq i})$. Strategy σ is *winning* for an epistemic objective $\varphi \in \mathcal{L}_{\text{EL}}$, if for every outcome π of σ there is some $i \in \mathbb{N}$ with $\pi_{\leq i} \models \varphi$.

Theorem 5 (Bozzelli et al., 2015) *The existence of a winning strategy for player 0 in an epistemic game \mathcal{G} for an epistemic objective φ of modal depth k can be decided in time k -exponential in $|\mathcal{G}|$ and $|\varphi|$.*

We show that the controller synthesis problem for propositional action models reduces to solving an epistemic game:

Proposition 6 *Given an instance $((\mathcal{M}, w_i), \mathcal{A}, \varphi)$ of the controller synthesis problem where \mathcal{A} is propositional, one can construct a game arena \mathcal{G} such that Controller wins in $((\mathcal{M}, w_i), \mathcal{A}, \varphi)$ iff Player 0 wins in \mathcal{G} for objective φ and $|\mathcal{G}| \leq |\mathcal{M}| + |\mathcal{A}| \times 2^{m+1}$, where m is the number of atomic propositions involved.*

Proof sketch. Let AP_u be the atomic propositions involved. The worlds of the game arena \mathcal{G} that we build are either worlds $w \in \mathcal{M}$ or tuples (α, v, i) where $\alpha \in \mathcal{A}$ represents the last action performed, $v \in 2^{AP_u}$ is the current valuation, and $i \in \{0, 1\}$ indicates whose turn it is to play: 0 for Controller and 1 for Environment. In an initial world w , Controller can choose an action $\alpha \in A_{\text{ctr}}$ such that $w \models \text{pre}(\alpha)$ and move to $(\alpha, V(w, \alpha), 1)$; in a world of the form (α, v, i) , if $i = 1$ (resp., $i = 0$), Environment (resp., Controller) chooses an action $\alpha' \in A_{\text{env}}$ (resp., $\alpha' \in A_{\text{ctr}}$) such that $v \models \text{pre}(\alpha')$, and moves to $(\alpha', V(w, \alpha), 1 - i)$. \square

Theorem 4 now follows from Theorem 5 and Proposition 6.

With the controller synthesis problem we enriched epistemic planning with an adversarial environment. Still, as in epistemic planning, the agents are mere observers. We now make a step further and make the agents players of the game.

5 Distributed Strategy Synthesis

In this section agents are no more passive, but instead they are players who choose themselves the actions that occur. The set Agt of agents is split into two teams Agt_{\exists} and Agt_{\forall} that play against each other, and we may say *players* instead of agents.

5.1 Setting up the Game

Unlike the external controller from the previous section, our players now have imperfect information. The fundamental feature of games with imperfect information is that when a player cannot distinguish between two different situations, a strategy for this player should prescribe the same action in both situations. All the additional complexity in solving games with imperfect information compared to the perfect information setting arises from this constraint. Such strategies are often called *uniform* or *observation-based* (see for instance [Reif, 1984; Van Benthem, 2001; Apt and Grädel, 2011]). Since games with imperfect information consider S5

models, i.e. where accessibility relations are equivalence relations, and it is unclear what uniform strategies mean in non-S5 models, we also assume from now on that all epistemic and action models are S5. We stress this assumption by writing \sim_a (resp. \sim_a^A) instead of R_a (resp. R_a^A).

We start from an initial pointed epistemic model \mathcal{M}, w , and an action model \mathcal{A} whose set of actions is partitioned into subsets $(A_a)_{a \in \text{Agt}}$ of actions for each player. The game we describe is turn-based. We use the variable turn ranging over Agt to represent whose turn it is to play. We require that for each $a \in A_a$, $\text{pre}(a)$ implies $(\text{turn}=a)$, and that postconditions for variable turn do not depend on the current world, but instead the next value of turn is completely determined by the action only. Moreover, in order to obtain a proper imperfect-information game, we demand the following hypotheses:

Hypotheses on \mathcal{M} and \mathcal{A}

- (H1) **The starting player is known:** there is a player a such that for all $u \in W$, $\mathcal{M}, u \models (\text{turn}=a)$;
- (H2) **The turn stays known:** for all actions α, α' and agent a , if $\alpha R_a \alpha'$, then α and α' assign the same value to turn.
- (H3) **Players know their available actions:** if agent a can execute α after history h , then she can also execute it after every history h' with $h \sim_a h'$.

All these hypotheses can be either enforced syntactically or checked in the different decidable cases we consider in the rest of this work (see the long version for detail).

We now define formally the notion of uniform strategies.

Definition 8 (Uniform strategy) *A strategy σ for player a is uniform if for every pair of histories h, h' where it is player a 's turn, $h \sim_a h'$ entails $\sigma(h) = \sigma(h')$.*

In the rest of this section, a strategy of a player in Agt_{\exists} is implicitly uniform. When one selects a strategy for each player in Agt_{\exists} , the result is called a *distributed strategy*, and an *outcome* of a distributed strategy is a play in which all players in Agt_{\exists} follow their prescribed strategy. A distributed strategy is *winning* for an objective formula φ if all its outcomes eventually satisfy φ .

5.2 The Distributed Strategy Synthesis Problem

We study the existence of a distributed strategy for players in Agt_{\exists} that ensures to reach a desired epistemic property.

Definition 9 (Distributed strategy synthesis problem)

- *Input:* a pointed epistemic model \mathcal{M}, w and an action model \mathcal{A} partitioned into $(A_a)_{a \in \text{Agt}}$ that satisfy hypotheses (H1)-(H3), and an objective formula $\varphi \in \mathcal{L}_{\text{EL}}$;
- *Output:* yes if there exists a winning distributed strategy for players in Agt_{\exists} ; no otherwise.

Unlike the controller synthesis problem, for propositional actions the distributed strategy synthesis problem is undecidable, already for a team of two players.

5.3 Undecidability for Two Existential Players

Because instances of the undecidable game TEAM DFA GAME [Coulombe and Lynch, 2018, Def. 1, p. 14:7] can be reduced to our distributed strategy synthesis problem, we get:

Theorem 7 *The distributed strategy synthesis problem is undecidable, already for a propositional action model and two existential players against one universal player.*

We now turn to decidable cases: games with imperfect information and epistemic objectives are known to be decidable either when actions are public [Belardinelli *et al.*, 2017b], or when information is hierarchical [Maubert and Murano, 2018]. We establish similar results in our setting.

5.4 The Case of Non-Expanding Action Models

Theorems 2 and 3 generalise to the distributed strategy synthesis problem. First, we inherit the lower bounds by letting $Agt_{\exists} = \{\text{Controller}\}$ and $Agt_{\forall} = \{\text{Environment}\}$, and by making them alternate turns. Second, the upper bounds are obtained by adapting the alternating algorithms for the upper bounds of Theorems 2 and 3. We need to ensure that existential choices of actions of an agent $a \in Agt_{\exists}$ lead to a *uniform* strategy: every time agent a picks an action α , we perform an extra universal choice over \sim_a -indistinguishable worlds, and continue executing the algorithm from these worlds.

Theorem 8 *For public announcements, the distributed strategy synthesis problem is PSPACE-complete.*

Theorem 9 *For public actions, the distributed strategy synthesis problem is EXPTIME-complete.*

We now turn to a decidable case for propositional actions.

5.5 The Case of Propositional Action Models and Hierarchical Information

We consider propositional action models, which may make the size of epistemic models grow unboundedly, but where the information of the different players is hierarchical, making it easier to synchronise the existential players' strategies.

According to Theorem 7, the distributed strategy synthesis problem is undecidable for propositional actions and a two-player team $Agt_{\exists} = \{a, b\}$ against team $Agt_{\forall} = \{\forall\}$. Observe that in the proof of Theorem 7, the information of players a and b is incomparable: in each round a only learns the first bit produced by \forall 's move, while b only learns the second bit. This cannot be simulated in games with so-called *hierarchical information*, a classic restriction to regain decidability in multi-player games of imperfect information [Peterson *et al.*, 2002; Pnueli and Rosner, 1990].

We say that an input of the distributed strategy synthesis problem $((\mathcal{M}, w_i), \mathcal{A}, \varphi)$ presents *hierarchical information* if the set of Agt_{\exists} can be totally ordered $(a_1 < \dots < a_n)$ so that $\sim_{a_i} \subseteq \sim_{a_{i+1}}$ and $\sim_{a_i}^A \subseteq \sim_{a_{i+1}}^A$, for each $1 \leq i < n$.

Theorem 10 *Distributed strategy synthesis with propositional actions and hierarchical information is decidable.*

We end the section by sketching the proof of Theorem 10. We start by introducing a multi-player variant of the epistemic game arenas from Definition 7.

Definition 10 *A multi-player epistemic game arena is given as $\mathcal{G} = (W, w_i, \Delta, (\sim_a)_{a \in Agt}, V)$ where $(W, (\sim_a)_{a \in Agt}, V)$ is an epistemic model, $W = \uplus_{a \in Agt} W_a$ is partitioned into positions of agents, w_i is an initial world and $\Delta \subseteq W \times W$ is a transition relation.*

Accessibility relations \sim_a are extended to histories, strategies of agent a are required to be uniform with respect to \sim_a , and the notions of outcomes, distributed strategies and winning distributed strategies are defined as before.

Theorem 10 is established by reducing the distributed strategy synthesis problem to a similar problem in multi-player epistemic games, known to be decidable:

Theorem 11 ([Maubert and Murano, 2018]) *Existence of winning distributed strategies in multi-player epistemic games with hierarchical information and epistemic temporal objectives is decidable.*

The reduction is very similar to the one in the proof of Proposition 6. The main difference is that we use variable turn instead of bit $i \in \{0, 1\}$ to define the positions of the different agents. The imperfect information of players is defined based on the accessibility relations in \mathcal{M} and \mathcal{A} .

6 Perspectives

We have incrementally extended the framework of epistemic planning to a game setting where players' actions are described by action models from DEL. We have established fine-grained results depending on the type of action models.

The problems we considered generalise main extensions of classical planning with game features: *conditional planning* with full (resp. partial) observability [Rintanen, 2004] is subsumed by our controller (resp. distributed strategy) synthesis problem. Also, *conformant planning* (partial information where the plan is a sequence of actions) corresponds to a particular case of our distributed strategy synthesis problem where $Agt_{\exists} = \{\exists\}$, $Agt_{\forall} = \{\forall\}$ are singletons, and \exists is blind, i.e. all actions in A_{\forall} are indistinguishable for her. Blindness and uniformity assumption make that the strategies of \exists can be seen as sequential plans.

Moreover, the decision problems we have considered go well beyond classical planning by addressing, e.g. distributed planning with cooperative or/and adversarial features. We are thus confronted in a DEL setting to issues usually met in game theory, as witnessed by the undecidability of the distributed strategy synthesis problem for rather simple action models (Theorem 7). However, the DEL perspective we propose offers a language to specify actions (preconditions, postconditions, and epistemic relations between actions) that may help identifying yet unknown decidable cases.

We note that our results should transfer to Game Description Language, equivalent to DEL [Engesser *et al.*, 2018].

One interesting extension of the unifying setting of DEL games would be to consider concurrent games, where players execute actions concurrently, but this will require to first generalise the product operation of DEL. Another direction would be to consider richer objectives such as ones expressed in epistemic temporal logic.

Our approach contributes to putting closer the field of multi-agent planning in AI with the field of multi-player games in formal methods. The setting of DEL games may be beneficial to both, allowing the transfer of powerful automata and game techniques from formal methods to epistemic planning, and bringing in multiplayer games new insights from the fine modelling offered by DEL.

References

- [Apt and Grädel, 2011] K.R. Apt and E. Grädel. *Lectures in game theory for computer scientists*. Cambridge UP, 2011.
- [Aucher *et al.*, 2014] G. Aucher, B. Maubert, and S. Pinchinat. Automata techniques for epistemic protocol synthesis. In *SR'14*, pages 97–103, 2014.
- [Baffier *et al.*, 2016] J.F. Baffier, M.K. Chiu, Y. Diez, M. Korman, V. Mitsou, A. Van Renssen, M. Roeloffzen, and Y. Uno. Hanabi is np-hard, even for cheaters who look at their cards. *arXiv preprint arXiv:1603.01911*, 2016.
- [Baltag *et al.*, 1998] A. Baltag, L. S Moss, and S. Solecki. The logic of public announcements, common knowledge, and private suspicions. In *TARK'98*, pages 43–56, 1998.
- [Belardinelli *et al.*, 2017a] F. Belardinelli, A. Lomuscio, A. Murano, and S. Rubin. Verification of broadcasting multi-agent systems against an epistemic strategy logic. In *IJCAI'17*, pages 91–97, 2017.
- [Belardinelli *et al.*, 2017b] F. Belardinelli, A. Lomuscio, A. Murano, and S. Rubin. Verification of multi-agent systems with imperfect information and public actions. In *AAMAS'17*, 2017.
- [Berwanger *et al.*, 2018] D. Berwanger, A. B. Mathew, and M. van den Bogaard. Hierarchical information and the synthesis of distributed strategies. *Acta Inf.*, 55(8), 2018.
- [Blumensath and Grädel, 2000] A. Blumensath and E. Grädel. Automatic structures. In *LICS'00*, pages 51–62, 2000.
- [Bolander and Andersen, 2011] T. Bolander and M.B. Andersen. Epistemic planning for single and multi-agent systems. *J. Appl. Non-Class. Log.*, 21(1):9–34, 2011.
- [Bolander *et al.*, 2015] T. Bolander, M.H. Jensen, and F. Schwarzenrüber. Complexity results in epistemic planning. In *IJCAI*, pages 2791–2797, 2015.
- [Bouyer, 2018] P. Bouyer. Games on graphs with a public signal monitoring. In *FOSSACS'18*. Springer, 2018.
- [Bozzelli *et al.*, 2015] L. Bozzelli, B. Maubert, and S. Pinchinat. Uniform strategies, rational relations and jumping automata. *Inf. Comput.*, 242:80–107, 2015.
- [Chandra and Stockmeyer, 1976] A.K. Chandra and L.J. Stockmeyer. Alternation. In *FOCS'76*, 1976.
- [Charrier, 2018] T. Charrier. *Theoretical complexity of reasoning in dynamic epistemic logic and symbolic approach*. PhD thesis, Université de Rennes 1, 2018.
- [Coulombe and Lynch, 2018] M. J. Coulombe and J. Lynch. Cooperating in video games? impossible! undecidability of team multiplayer games. In *FUN'18*, pages 14:1–14:16, 2018.
- [Douéneau-Tabot *et al.*, 2018] G. Douéneau-Tabot, S. Pinchinat, and F. Schwarzenrüber. Chain-monadic second order logic over regular automatic trees and epistemic planning synthesis. In *AiML'18*, 2018.
- [Engesser *et al.*, 2017] T. Engesser, T. Bolander, R. Mattmüller, and B. Nebel. Cooperative epistemic multi-agent planning for implicit coordination. In *M4M*, pages 75–90, 2017.
- [Engesser *et al.*, 2018] T. Engesser, R. Mattmüller, B. Nebel, and M. Thielscher. Game description language and dynamic epistemic logic compared. In *IJCAI'18*, pages 1795–1802, 2018.
- [Fagin *et al.*, 1995] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.
- [Geffner and Bonet, 2013] Hector Geffner and Blai Bonet. A concise introduction to models and methods for automated planning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 8(1):1–141, 2013.
- [Harel and Pnueli, 1985] D. Harel and A. Pnueli. On the development of reactive systems. In *Logics and models of concurrent systems*, pages 477–498. Springer, 1985.
- [Lê Cong *et al.*, 2018] S. Lê Cong, S. Pinchinat, and F. Schwarzenrüber. Small undecidable problems in epistemic planning. In *IJCAI'18*, 2018.
- [Lima, 2014] T. De Lima. Alternating-time temporal dynamic epistemic logic. *J. Log. Comput.*, 24(6):1145–1178, 2014.
- [Littman *et al.*, 1998] M.L. Littman, J. Goldsmith, and M. Mundhenk. The computational complexity of probabilistic planning. *J. Artif. Intell. Res.*, 9:1–36, 1998.
- [Matros, 2018] A. Matros. Lloyd shapley and chess with imperfect information. *Games and Economic Behavior*, 108:600–613, 2018.
- [Maubert and Murano, 2018] B. Maubert and A. Murano. Reasoning about knowledge and strategies under hierarchical information. In *KR'18*, pages 530–540, 2018.
- [Maubert *et al.*, 2019] B. Maubert, S. Pinchinat, and F. Schwarzenrüber. Reachability games in dynamic epistemic logic. Research Report 1905.12422, arXiv, May 2019.
- [Maubert, 2014] B. Maubert. *Logical foundations of games with imperfect information : uniform strategies*. PhD thesis, University of Rennes 1, France, 2014.
- [Peterson and Reif, 1979] G.L. Peterson and J.H. Reif. Multiple-person alternation. In *FOCS'79*, 1979.
- [Peterson *et al.*, 2002] G. Peterson, J. Reif, and S. Azhar. Decision algorithms for multiplayer noncooperative games of incomplete information. *Computers & Mathematics with Applications*, 43(1):179–206, 2002.
- [Pnueli and Rosner, 1990] A. Pnueli and R. Rosner. Distributed reactive systems are hard to synthesize. In *FOCS'90*, pages 746–757, 1990.
- [Ramanujam and Simon, 2010] R. Ramanujam and S. Simon. A communication based model for games of imperfect information. In *CONCUR'10*, pages 509–523, 2010.
- [Reif, 1984] John H. Reif. The complexity of two-player games of incomplete information. *J. Comput. Syst. Sci.*, 29(2):274–301, 1984.
- [Rintanen, 2004] J. Rintanen. Complexity of planning with partial observability. In *ICAPS'04*, pages 345–354, 2004.
- [Van Benthem, 2001] J. Van Benthem. Games in dynamic-epistemic logic. *Bulletin of Economic Research*, 53(4):219–248, 2001.
- [van Ditmarsch *et al.*, 2008] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, Dordrecht, 2008.
- [Yu *et al.*, 2013] Q. Yu, X. Wen, and Y. Liu. Multi-agent epistemic explanatory diagnosis via reasoning about actions. In *IJCAI'13*, 2013.