Goal Recognition as Planning with Unknown Domain Models

Diego Aineto[1] , **Sergio Jiménez**[1] , **Eva Onaindia**[1] and , **Miquel Ramírez**[2]

[1]Departamento de Sistemas Informáticos y Computación. Universitat Politècnica de València. Valencia, Spain
[2]School of Computing and Information Systems. The University of Melbourne. Melbourne, Victoria. Australia

{dieaigar,serjice,onaindia}@dsic.upv.es, miquel.ramirez@unimelb.edu.au

Abstract

This paper shows how to relax a strong assumption of the *plan recognition as planning* approach for *goal recognition* that is a priori having an action model of the observed agents. The paper introduces a novel formulation for classical planning in a setting where no action model is given (instead, only the state variables and the action headers are known) and it shows that this formulation neatly fits with the popular *plan recognition as planning* approach for *goal recognition*. The empirical evaluation evidences that this novel formulation allows to solve standard goal recognition benchmarks without *a priori* knowing the action model of the observed agents and using an off-the-shelf classical planner.

1 Introduction

Goal recognition is a particular classification task in which each class represents a different goal and the classification examples are observations of agents acting to achieve one of that goals. Despite there exists a wide range of different approaches for *goal recognition*, *plan recognition as planning* [Ramírez and Geffner, 2009; Ramírez, 2012] is one of the most appealing and it is currently at the core of various activity recognition tasks such as, *goal recognition design* [Keren *et al.*, 2014], *deceptive planning* [Masters and Sardina, 2017], *planning for transparency* [MacNally *et al.*, 2018] or *counter-planning* [Pozanco *et al.*, 2018].

Plan recognition as planning leverages the action model of the observed agents and an off-the-shelf classical planner to compute the most likely goal of that agents. In this paper we show how to relax the assumption of *knowing the action model of the observed agents*, which frequently becomes a too strong assumption when using *plan recognition as planning* at real-world applications. In particular, the paper introduces a novel formulation for classical planning in a setting where no action model is given (instead, only the state variables and the action headers are given) and it shows that this formulation neatly fits with the *plan recognition as planning* approach. The empirical evaluation evidences that this novel formulation allows to solve standard goal recognition benchmarks without *a priori* knowing the action model of the observed agents and using an off-the-shelf classical planner.

2 Background

This section formalizes the *planning model* we follow, the kind of *observations* that are given as input to the *goal recognition* task, and the *plan recognition as planning* approach for *goal recognition*.

2.1 Classical planning with conditional effects

Let F be the set of propositional state variables (*fluents*) describing a state. A *literal* l is a valuation of a fluent $f \in F$; i.e. either $l = f$ or $l = \neg f$. A set of literals L represents a partial assignment of values to fluents (without loss of generality, we will assume that L does not contain conflicting values). Given L, let $\neg L = \{\neg l : l \in L\}$ be its complement. We use $\mathcal{L}(F)$ to denote the set of all literal sets on F; i.e. all partial assignments of values to fluents. A *state* s is a full assignment of values to fluents; $|s| = |F|$.

A classical planning action $a \in A$ has: a precondition $\mathsf{pre}(a) \in \mathcal{L}(F)$, a set of effects $\mathsf{eff}(a) \in \mathcal{L}(F)$, and a positive action cost $cost(a)$. The semantics of actions $a \in A$ is specified with two functions: $\rho(s, a)$ denotes whether action a is *applicable* in a state s and $\theta(s, a)$ denotes the *successor state* that results of applying action a in a state s. Then, $\rho(s, a)$ holds iff $\mathsf{pre}(a) \subseteq s$, i.e. if its precondition holds in s. The result of executing an applicable action $a \in A$ in a state s is a new state $\theta(s, a) = (s \setminus \neg\mathsf{eff}(a)) \cup \mathsf{eff}(a)$. Subtracting the complement of $\mathsf{eff}(a)$ from s ensures that $\theta(s, a)$ remains a well-defined state. The subset of action effects that assign a positive value to a state fluent is called *positive effects* and denoted by $\mathsf{eff}^+(a) \in \mathsf{eff}(a)$ while $\mathsf{eff}^-(a) \in \mathsf{eff}(a)$ denotes the *negative effects* of an action $a \in A$.

A *classical planning problem* is a tuple $P = \langle F, A, I, G\rangle$, where I is the initial state and $G \in \mathcal{L}(F)$ is the set of goal conditions over the state variables. A *plan* π is an action sequence $\pi = \langle a_1, \ldots, a_n\rangle$, with $|\pi| = n$ denoting its *plan length* and $cost(\pi) = \sum_{a\in\pi} cost(a)$ its *plan cost*. The execution of π on the initial state I of P induces a *trajectory* $\tau(\pi, s_0) = \langle s_0, a_1, s_1, \ldots, a_n, s_n\rangle$ such that $s_0 = I$ and, for each $1 \leq i \leq n$, it holds $\rho(s_{i-1}, a_i)$ and $s_i = \theta(s_{i-1}, a_i)$. A plan π solves P iff the induced *trajectory* $\tau(\pi, s_0)$ reaches a final state $G \subseteq s_n$, where all goal conditions are met. A solution plan is *optimal* iff its cost is minimal.

An *action with conditional effects* $a_c \in A$ is defined as a set of preconditions $\mathsf{pre}(a_c) \in \mathcal{L}(F)$ and a set of *conditional effects* $\mathsf{cond}(a_c)$. Each conditional effect $C \rhd E \in \mathsf{cond}(a_c)$ is composed of two sets of literals: $C \in \mathcal{L}(F)$, the *condition*, and $E \in \mathcal{L}(F)$, the *effect*. An action a_c is applicable in a state s if $\rho(s, a_c)$ is true, and the result of applying action a_c in state s is $\theta(s, a_c) = \{s \setminus \neg\mathsf{eff}_c(s,a) \cup \mathsf{eff}_c(s,a)\}$ where $\mathsf{eff}_c(s,a)$ are the *triggered effects* resulting from the action application (conditional effects whose conditions hold in s):

$$\mathsf{eff}_c(s,a) = \bigcup_{C \rhd E \in \mathsf{cond}(a_c), C \subseteq s} E,$$

2.2 The observation model

Given a planning problem $P = \langle F, A, I, G\rangle$, a plan π and a trajectory $\tau(\pi, P)$, we define the *observation of the trajectory* as an interleaved combination of actions and states that represents the observation from the execution of π in P. Formally, $\mathcal{O}(\tau) = \langle s_0^o, a_1^o, s_1^o \ldots, a_l^o, s_m^o\rangle$, $s_0^o = I$, and:

- The **observed actions** are consistent with π, which means that $\langle a_1^o, \ldots, a_l^o\rangle$ is a sub-sequence of π. Specifically, the number of observed actions, l, can range from 0 (fully unobservable action sequence) to $|\pi|$ (fully observable action sequence).
- The **observed states** $\langle s_0^o, s_1^o, \ldots, s_m^o\rangle$ is a sequence of possibly *partially observable states*, except for the initial state s_0^o, which is fully observable. A partially observable state s_i^o is one in which $|s_i^o| < |F|$; i.e., a state in which at least a fluent of F is not observable. Note that this definition also comprises the case $|s_i^o| = 0$, when the state is fully unobservable. Whatever the sequence of observed states of $\mathcal{O}(\tau)$ is, it must be consistent with the sequence of states of $\tau(\pi, P)$, meaning that $\forall i, s_i^o \subseteq s_i$. In practice, the number of observed states, m, range from 1 (the initial state, at least), to $|\pi|+1$, and the observed intermediate states will comprise a number of fluents between $[1, |F|]$.

We assume a bijective monotone mapping between actions/states of trajectories and observations [Ramírez and Geffner, 2009], thus also granting the inverse consistency relationship (the trajectory is a superset of the observation). Therefore, transiting between two consecutive observed states in $\mathcal{O}(\tau)$ may require the execution of more than a single action ($\theta(s_i^o, \langle a_1, \ldots, a_k\rangle) = s_{i+1}^o$, where $k \geq 1$ is unknown but finite. In other words, having $\mathcal{O}(\tau)$ does not imply knowing the actual length of π.

2.3 Goal recognition with classical planning

Goal recognition is a particular classification task in which each class represents a different goal $g \in G[\cdot]$ and there is a single classification example $\mathcal{O}(\tau)$ that represents the observation of agents acting to achieve one of the input goals in $g \in G[\cdot]$. Following the *naive Bayes classifier*, the *solution* to the *goal recognition* task is the subset of goals in $G[\cdot]$ that maximizes this expression.

$$argmax_{g \in G[\cdot]} P(\mathcal{O}|g)P(g). \tag{1}$$

The *plan recognition as planning* approach shows how to compute estimates of the $P(\mathcal{O}|g)$ likelihood leveraging the action model of the observed agent and an off-the-shelf classical planner [Ramírez, 2012]. Given a *classical planning problem* $P = \langle F, A, I, G[\cdot]\rangle$, where $G[\cdot]$ represents the set of possible goals $P(\mathcal{O}|g)$ is estimated computing, for each goal $g \in G[\cdot]$, the cost difference of the solution plans to these two different classical planning problems:

- $P_g^\top$, the classical planning problem built constraining $P = \langle F, A, I, g\rangle$ to achieve the particular goal $g \in G[\cdot]$ through a plan $\pi^\top$ *consistent* with the input observation $\mathcal{O}(\tau)$.
- $P_g^\perp$, the classical planning problem that constrains the solutions of $P = \langle F, A, I, g\rangle$ to plans $\pi^\perp$ that achieve $g \in G[\cdot]$ but are *inconsistent* with $\mathcal{O}(\tau)$.

The higher the value of this cost difference $cost(\pi^\top) - cost(\pi^\perp)$, the higher probability of aiming to achieve the $g \in G[\cdot]$ goal. With this regard, *plan recognition as planning* uses the *sigmoid function* to map the previous cost difference into a likelihood:

$$P(\mathcal{O}|g) = \frac{1}{1 + e^{-\beta(cost(\pi^\top) - cost(\pi^\perp))}} \tag{2}$$

This expression is derived from the assumption that while the observed agent is not perfectly rational, he is more likely to follow cheaper plans, according to a *Boltzmann* distribution. The larger the value of β, the more rational the agent, and the less likely that he will follow suboptimal plans. Recent works show that estimates of the $P(\mathcal{O}|g)$ likelihood can be faster computed using relaxations of the classical planning tasks [Pereira *et al.*, 2017].

The work on *plan recognition as planning* usually assumes an observation model that is referred only to logs of executed actions [Ramírez and Geffner, 2009]. However, the approach applies to more expressive observation models that consider also state observations, like the one defined above, with a trivial three-fold extension:

- One fluent $\{validated_j\}_{0 \leq j \leq m}$ to point at every $s_j \in \mathcal{O}(\tau)$ state observation.
- Adding $validated_m$ to every possible goal $g \in G[\cdot]$ to constraint the solution plans $\pi^\top$ to be consistent with all the state observations.
- One $\mathsf{validate_j}$ action to constraint $\pi^\top$ to be consistent with the $s_j \in \mathcal{O}(\tau)$ input state observation, $(1 \leq j \leq m)$.

$$\begin{aligned}\mathsf{pre}(\mathsf{validate_j}) =& s_j \cup \{validated_{j-1}\},\\ \mathsf{cond}(\mathsf{validate_j}) =& \{\emptyset\} \rhd \{\neg validated_{j-1}, validated_j\}.\end{aligned}$$

3 Planning with unknown domain models

This section introduces a novel formulation for classical planning in a setting where no action model is given. This setting of classical planning is very related to the learning of action models for planning [Stern and Juba, 2017]. In particular it can be considered an extreme learning scenario when a model has to be learned from a single learning example that contains only two state observations, an initial state and the goals.

A *classical planning with unknown domain models* is a tuple $P = \langle F, A[\cdot], I, G\rangle$, where $A[\cdot]$ is a set of actions s.t., the semantics of each action $a \in A[\cdot]$ is unknown (i.e. the functions ρ and/or θ of a are undefined).

A solution to this task is a sequence of actions $\pi = \langle a_1, \ldots, a_n\rangle$ whose execution induces a trajectory $\tau(\pi, I) = \langle s_0, a_1, s_1, \ldots, a_n, s_n\rangle$ such that $s_0 = I$ and *there exists* at least one possible action model (e.g. one possible definition of the ρ and θ functions within the given state variables) that satisfies $\rho(s_{i-1}, a_i)$ and $s_i = \theta(s_{i-1}, a_i)$, for each $1 \leq i \leq n$, and such that the reached final state met the goal conditions, $G \subseteq s_n$.

Next we show that the space of possible STRIPS action models can be encoded as a set of propositional variables and a set of constraints over those variables. Then, we show how to exploit this encoding to solve $P = \langle F, A[\cdot], I, G\rangle$ problems with an off-the-shelf classical planner as well as the properties of this approach.

3.1 A propositional encoding for the space of STRIPS action models

A STRIPS *action schema* ξ is defined by four lists: A list of *parameters* $pars(\xi)$, and three list of predicates (namely $pre(\xi)$, $del(\xi)$ and $add(\xi)$) that shape the kind of fluents that can appear in the *preconditions*, *negative effects* and *positive effects* of the actions induced from that schema. Let be Ψ the set of *predicates* that shape the propositional state variables F, and a list of *parameters* $pars(\xi)$. The set of elements that can appear in $pre(\xi)$, $del(\xi)$ and $add(\xi)$ of the STRIPS action schema ξ is given by FOL interpretations of Ψ over the parameters $pars(\xi)$ and is denoted as $\mathcal{I}_{\Psi,\xi}$.

For instance, in the *blocksworld* the $\mathcal{I}_{\Psi,\xi}$ set contains only five elements for a `pickup(`v_1`)` schemata, $\mathcal{I}_{\Psi,pickup}$={`handempty, holding(`$v_1$`), clear(`$v_1$`), ontable(`$v_1$`), on(`$v_1$`,`$v_1$`)`} while it contains eleven elements for a `stack(`v_1`,`v_2`)` schemata, $\mathcal{I}_{\Psi,stack}$={`handempty, holding(`v_1`), holding(`v_2`), clear(`v_1`), clear(`v_2`), ontable(`v_1`), ontable(`v_2`), on(`v_1`,`v_1`), on(`v_1`,`v_2`), on(`v_2`,`v_1`), on(`v_2`,`v_2`)`}.

Despite any element of $\mathcal{I}_{\Psi,\xi}$ can *a priori* appear in the $pre(\xi)$, $del(\xi)$ and $add(\xi)$ of schema ξ, the actual space of possible STRIPS schemata is bounded by constraints of three kinds:

1. *Syntactic constraints*. STRIPS constraints require $del(\xi) \subseteq pre(\xi)$, $del(\xi) \cap add(\xi) = \emptyset$ and $pre(\xi) \cap add(\xi) = \emptyset$. Considering exclusively these syntactic constraints, the size of the space of possible STRIPS schemata is given by $2^{2\times|\mathcal{I}_{\Psi,\xi}|}$. *Typing constraints* are also of this kind [McDermott *et al.*, 1998].
2. *Domain-specific constraints*. One can introduce domain-specific knowledge to constrain further the space of possible schemata. For instance, in the *blocksworld* one can argue that `on(`v_1`,`v_1`)` and `on(`v_2`,`v_2`)` will not appear in the $pre(\xi)$, $del(\xi)$ and $add(\xi)$ lists of an action schema ξ because, in this specific domain, a block cannot be on top of itself. *State invariants* are also constraints of this kind [Fox and Long, 1998].
3. *Observation constraints*. An observations $\mathcal{O}(\tau)$ depicts *semantic knowledge* that constraints further the space of possible action schemata.

```
(:action stack
   :parameters (?v1 ?v2)
   :precondition (and (holding ?v1) (clear ?v2))
   :effect (and (not (holding ?v1)) (not (clear ?v2))
                (clear ?v1) (handempty) (on ?v1 ?v2)))


(pre_holding_v1_stack) (pre_clear_v2_stack)
(eff_holding_v1_stack) (eff_clear_v2_stack)
(eff_clear_v1_stack) (eff_handempty_stack) (eff_on_v1_v2_stack)
```

Figure 1: PDDL encoding of the `stack(?v1,?v2)` schema and our propositional representation for this same schema.

In this work we introduce a propositional encoding of the *preconditions*, *negative*, and *positive* effects of a STRIPS action schema ξ using only fluents of two kinds `pre_e_`ξ and `eff_e_`ξ (where $e \in \mathcal{I}_{\Psi,\xi}$). This encoding exploits the syntactic constraints of STRIPS so is more compact that the one previously proposed by Aineto *et al.* 2018. In more detail, if `pre_e_`ξ and `eff_e_`ξ holds it means that $e \in \mathcal{I}_{\Psi,\xi}$ is a *negative effect* in ξ while if *pre_e_*ξ does not hold but `eff_e_`ξ holds, it means that $e \in \mathcal{I}_{\Psi,\xi}$ is a *positive effect* in ξ. Figure 1 shows the PDDL encoding of the `stack(?v1,?v2)` schema and our propositional representation for this same schema with `pre_e_stack` and `eff_e_stack` fluents ($e \in \mathcal{I}_{\Psi,stack}$).

3.2 A classical planning compilation for planning with unknown domain models

Now we show how we adapt the classical planning compilation [Aineto *et al.*, 2018] to our propositional encoding. In more detail, given a classical planning problem with unknown domain models $P = \langle F, A[\cdot], I, G\rangle$ we create another classical planning problem $P' = \langle F', A', I, G\rangle$ such that:

- F' extends F with the fluents for the propositional encoding of the corresponding space of STRIPS action models. This is a set of fluents of the type $\{pre\_e\_\xi, eff\_e\_\xi\}_{\forall e \in \mathcal{I}_{\Psi,\xi}}$ such that $e \in \mathcal{I}_{\Psi,\xi}$ is a single element from the set of FOL interpretations of predicates Ψ over the corresponding parameters $pars(\xi)$.
- A' replaces the actions in A with two new types of actions.
  1. Actions for *inserting* a component (precondition, positive effect or negative effect) in $\xi \in \mathcal{M}$ following the syntactic constraints of STRIPS models.
     - Actions which support the addition of a *precondition* $p \in \Psi_\xi$ to the action model $\xi \in \mathcal{M}$. A precondition p is inserted in ξ when neither pre_p, eff_p exist in ξ.

$$\mathsf{pre}(\mathsf{insertPre}_{\mathsf{p},\xi}) = \{\neg pre_p(\xi), \neg eff_p(\xi)\},$$
$$\mathsf{cond}(\mathsf{insertPre}_{\mathsf{p},\xi}) = \{\emptyset\} \rhd \{pre_p(\xi)\}.$$

– Actions which support the addition of a *negative* or *positive* effect $p \in \Psi_\xi$ to the action model $\xi \in \mathcal{M}$.

$$\begin{aligned}\mathsf{pre}(\mathsf{insertEff}_{\mathsf{p},\xi}) =& \{\neg eff_p(\xi)\},\\ \mathsf{cond}(\mathsf{insertEff}_{\mathsf{p},\xi}) =& \{\emptyset\} \rhd \{eff_p(\xi)\}.\end{aligned}$$

2. Actions for *applying* the action models $\xi \in \mathcal{M}$ built by the insert actions and bounded to objects $\omega \subseteq \Omega^{ar(\xi)}$. Since action headers are known, the variables $pars(\xi)$ are bounded to the objects in ω that appear in the same position.

$$\begin{aligned}\mathsf{pre}(\mathsf{apply}_{\xi,\omega}) =& \{pre_p(\xi) \implies p(\omega)\}_{\forall p \in \Psi_\xi},\\ \mathsf{cond}(\mathsf{apply}_{\xi,\omega}) =& \{pre_p(\xi) \wedge eff_p(\xi)\} \rhd \{\neg p(\omega)\}_{\forall p \in \Psi_\xi},\\ & \{\neg pre_p(\xi) \wedge eff_p(\xi)\} \rhd \{p(\omega)\}_{\forall p \in \Psi_\xi}.\end{aligned}$$

The dynamics of the actions for *applying* an action model $\xi \in \mathcal{M}$ is determined by the values of the model the $\{pre\_e\_\xi, eff\_e\_\xi\}_{\forall e \in \mathcal{I}_{\Psi,\xi}}$ fluents in the current state. For instance, executing the action `(apply_stack blockB blockA)` in a state s implies activating the preconditions and effects of `(apply_stack)` according to the values of the model fluents in s. This means that if the current state s holds `{(pre_stack_holding_v1),(pre_stack_clear_v2)}` $\subset s$, then it must be checked that positive literals `(holding blockB)` and `(clear blockA)` hold in s. Otherwise, a different set of precondition literals will be checked. The same applies to the conditional effects, generating the corresponding literals according to the values of the model fluents of s. Note that executing `(apply_stack blockB blockA)`, will add the literals `(on blockB blockA)`,`(clear blockB)`,`(not(clear blockA))`,`(handempty)` and `(not(clear blockB))` to the successor state if `stack` has been correctly programmed by the insert actions.

```
00 : (insert_pre_stack_holding_v1)
01 : (insert_pre_stack_clear_v2)
02 : (insert_eff_stack_clear_v1)
03 : (insert_eff_stack_clear_v2)
04 : (insert_eff_stack_handempty)
05 : (insert_eff_stack_holding_v1)
06 : (insert_eff_stack_on_v1_v2)
07 : (apply_stack blockA blockB)
```

Figure 2: Plan computed when solving the classical planning problem output by our compilation for solving a classical planning with unknown domain models.

The plan of Figure 2 shows a solution plan computed when solving the classical planning problem output by our compilation for solving a classical planning with unknown domain models. In the initial state of that problem the robot is holding the blockA while the blockB is clear and the problem goal is having blockA on blockB. The plan shows the insert actions for the action model `stack` (steps $00 - 01$ insert the preconditions of the `stack` model, steps $02 - 06$ insert the action model effects), and step 07 is the plan postfix that applies the action models to achieve the goals. Note that an equivalent solution could be found by inserting the same preconditions and effects into the *unstack* action model and then applying action (unstack blockA blockB) instead of (stack blockA blockB).

3.3 Compilation properties

Now we present the properties of the compilation scheme.

Lemma 1. *Soundness. Any classical plan π' that solves P' produces a solution to the classical planning problem with unknown domain models $P = \langle F, A[\cdot], I, G\rangle$.*

Proof. Once a given precondition or effect is inserted into an action model it can never be removed back. In addition, P' is only solvable iff goals G hold at the last state reached by π'. In the compiled problem the value of F, the fluents of the original problem, can exclusively be modified via $\mathsf{apply}_{\xi,\omega}$ actions. The set of goals G can only be achieved executing an applicable sequence of $\mathsf{apply}_{\xi,\omega}$ actions that, starting in the corresponding initial state reach a state $G \subseteq s_n$. This means that the action model used by the $\mathsf{apply}_{\xi,\omega}$ actions has to be consistent with the traversed intermediate states. We know that this is true by the definition of the $\mathsf{apply}_{\xi,\omega}$ so hence, the sub-sequence of $\mathsf{apply}_{\xi,\omega}$ appearing in π' to solve P' is a solution plan to $P = \langle F, A[\cdot], I, G\rangle$. □

Lemma 2. *Completeness. Any plan π that solves $P = \langle F, A[\cdot], I, G\rangle$ is computable solving the corresponding classical planning task P'.*

Proof. $\mathcal{I}_{\Psi,\xi}$ fully captures the set of elements that can appear in an action model ξ using the set of predicates Ψ. The compilation does not discard any possible action model definable within $\mathcal{I}_{\Psi,\xi}$. This means that for every plan π that solves $P = \langle F, A[\cdot], I, G\rangle$, we can build a plan π' selecting the appropriate actions for inserting precondition and effects to the corresponding action model and then selecting the $\mathsf{apply}_{\xi,\omega}$ actions that allows to transform the initial state I into a state $G \subseteq s_n$. □

The bias of the initially empty action model

Given that at the initial state of the classical planning compilation any $\{pre\_e\_\xi, eff\_e\_\xi\}_{\forall e \in \mathcal{I}_{\Psi,\xi}}$ fluent is false, our compilation introduces a bias to solutions of the $P = \langle F, A[\cdot], I, G\rangle$ classical planning task.

This bias might be eliminated defining a cost landscape where any actions for *inserting* a precondition or an effect into a given action model has a zero cost while the remaining actions, $\mathsf{apply}_{\xi,\omega}$ actions has unit cost. In practice, since classical planners are not proficiency optimizing the cost of solutions with cost landscapes of this kind, we use a different approach that disregards the cost of the actions that insert a precondition/effect to an action model.

Our approach is to use SAT-based planning because it can apply all the required actions for inserting preconditions in a single planning step (in parallel), because these actions do not interact. Actions for programming action effects are also applied in a single planning step so the plan horizon for programming any action model is always bound to 2, which in addition significantly reduces the planning horizon. The SAT-based planning approach is also convenient because its ability to deal with planning problems populated with dead-ends and because the symmetries in the insertion of preconditions/effects into the action model do not affect the performance of SAT-based planning.

Compilation size

The size of the classical planning task P' output by the compilation approach depends on the arity of the *predicates* Ψ, that shape the propositional state variables F, and the arity of the action headers, given by $A[\cdot]$. The larger the arity, the larger the size of the $\mathcal{I}_{\Psi,\xi}$ sets. This is the term that dominates the compilation size because it defines the $pre_p(\xi)/del_p(\xi)/add_p(\xi)$ fluents, the corresponding set of *insert* actions, and the number of conditional effects in the $\mathsf{apply}_{\xi,\omega}$ actions.

Note that for planning models that allow object typing, types can be used straightforward to constrain the FOL interpretations of Ψ over the parameters $pars(\xi)$ significantly reducing the size of the classical planning task output by the compilation.

4 Goal recognition as planning with unknown domain models

We define the task of *goal recognition with unknown domain models* as a $\langle P, \mathcal{O}(\tau)\rangle$ pair, where:

- $P = \langle F, A[\cdot], I, G[\cdot]\rangle$ is a classical planning problem where $G[\cdot]$ is the set of possible goals and $A[\cdot]$ is a set of actions s.t., for each $a \in A[\cdot]$, the semantics of a is unknown (i.e. the functions ρ and/or θ of a are undefined).
- $\mathcal{O}(\tau)$ is an observation of a trajectory $\tau(\pi, P)$ produced by the execution of an unknown plan π that reaches a goal $g \in G[\cdot]$ starting from the given initial state.

The *solution* to the *goal recognition with unknown domain models* task is again the subset of goals in $G[\cdot]$ that maximizes expression (1).

4.1 Estimating $P(\mathcal{O}|g)$ with classical planning

Now we are ready to compute an estimate of the target probability distribution $P(g|\mathcal{O})$ over the possible goals $g \in G[\cdot]$ given an observation $\mathcal{O}(\tau)$:

1. For each goal $g \in G[\cdot]$,
   (a) Build the classical planning problem $P_g^\top$, that constrains the solutions to $P = \langle F, A[\cdot], s_0, g\rangle$ to plans $\pi^\top$ *consistent* with the input observation $\mathcal{O}(\tau)$. Note that s_0 is the initial state in the given observation $\mathcal{O}(\tau)$.
   (b) Solve $P_g^\top$ using the proposed compilation and a classical planner. Extract from this solution $cost(\pi^\top)$ (by counting the number of $\mathsf{apply}_{\xi,\omega}$ actions in the solution) and the action model that was determined by the insert actions to achieve the goals.
   (c) Build the classical planning problem $P_g^\perp$ that constrains $P = \langle F, A, s_0, g\rangle$ to achieve $g \in G[\cdot]$ through a plan $\pi^\perp$ *inconsistent* with $\mathcal{O}(\tau)$ and that uses the action model A determined by the solution to the proposed compilation.
   (d) Compute the cost difference $\Delta(cost(\pi_\top), cost(\pi_\perp))$ and plug this cost difference into equation (2) to get the $P(g|\mathcal{O})$ likelihoods.
2. Plug the likelihoods into the Bayes rule from which the goal posterior probabilities are obtained. In this case the $P\mathcal{O}(\tau)$ probabilities are obtained by normalization (goal probabilities must add up to 1 when summed over all possible goals).

5 Evaluation

6 Related Work

The problem of *classical planning with unknown domain models* has been previously addressed for SAS^+ action models [Stern and Juba, 2017]. In this work we formulate this task for STRIPS action models and evidence its relevance for addressing *goal recognition* tasks when the action model of the observed agent is not available (which typically is a too strong assumption for many real-world applications).

The paper also evidenced that *goal recognition*, when the domain Model is unknown, is very related to the task of learning a classical planning action model. With this regard the classical planning compilation for learning STRIPS action model is very suitable because it allows to produce a STRIPS action model from minimal input knowledge (and initial state, goals pair), and to refine this model if more knowledge is available (e.g. observation constraints) [Aineto *et al.*, 2018]. Most of the existing approaches for learning action models aim maximizing the statistical consistency of the learned model with respect to observations so require large amounts of input knowledge and do not produce model that are guaranteed to be *logically consistent* with the given input knowledge.

Planning with an unkown model is also related to *goal recognition design* [Keren *et al.*, 2014] where they compute the modification of the initial state that allow an efficient online goal recognition. Our approach is related since we are encoding the space of propositional schemas planning state variables, we encode at the initial state the *empty* actoin model with no preconditions and effects and provide actions to modify this model to enable the *goal recognition* process.

7 Conclusions

References

[Aineto *et al.*, 2018] Diego Aineto, Sergio Jiménez, and Eva Onaindia. Learning STRIPS action models with classical planning. In *International Conference on Automated Planning and Scheduling, (ICAPS-18)*, pages 399–407. AAAI Press, 2018.

[Fox and Long, 1998] Maria Fox and Derek Long. The automatic inference of state invariants in tim. *Journal of Artificial Intelligence Research*, 9:367–421, 1998.

[Keren *et al.*, 2014] Sarah Keren, Avigdor Gal, and Erez Karpas. Goal recognition design. In *International Conference on Automated Planning and Scheduling, (ICAPS-14)*, pages 154–162, 2014.

[MacNally *et al.*, 2018] Aleck M MacNally, Nir Lipovetzky, Miquel Ramirez, and Adrian R Pearce. Action selection for transparent planning. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1327–1335. International Foundation for Autonomous Agents and Multiagent Systems, 2018.

[Masters and Sardina, 2017] Peta Masters and Sebastian Sardina. Deceptive path-planning. In *IJCAI 2017*, pages 4368–4375. AAAI Press, 2017.

[McDermott *et al.*, 1998] Drew McDermott, Malik Ghallab, Adele Howe, Craig Knoblock, Ashwin Ram, Manuela Veloso, Daniel Weld, and David Wilkins. PDDL – The Planning Domain Definition Language, 1998.

[Pereira *et al.*, 2017] Ramon Fraga Pereira, Nir Oren, and Felipe Meneguzzi. Landmark-based heuristics for goal recognition. In *Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*. AAAI Press, 2017.

[Pozanco *et al.*, 2018] Alberto Pozanco, Yolanda E.-Martín, Susana Fernández, and Daniel Borrajo. Counterplanning using goal recognition and landmarks. In *International Joint Conference on Artificial Intelligence, (IJCAI-18)*, pages 4808–4814, 2018.

[Ramírez and Geffner, 2009] Miquel Ramírez and Hector Geffner. Plan recognition as planning. In *International Joint conference on Artifical Intelligence, (IJCAI-09)*, pages 1778–1783. AAAI Press, 2009.

[Ramírez, 2012] Miquel Ramírez. *Plan recognition as planning*. PhD thesis, Universitat Pompeu Fabra, 2012.

[Stern and Juba, 2017] Roni Stern and Brendan Juba. Efficient, safe, and probably approximately complete learning of action models. In *International Joint Conference on Artificial Intelligence, (IJCAI-17)*, pages 4405–4411, 2017.