Safe Learning of PDDL Domains with Conditional Effects - Supplementary Material

Primary Keywords: *(2) Learning; (8) Knowledge Representation/Engineering*

Domain Information

In Table 1 we present the information on the domains containing conditional effects available at https://github.com/AI-Planning/classical-domains. For each domain we describe the properties observed in the domain, whether the used solvers, i.e. Fast Downward and Fast Forward were able to solve the dataset problems with the *original* PDDL domain, and whether we used the domain in our experiments. The properties shown in the domains are as follows:

- Disjunctive antecedents - i.e., whether a conditional effect can appear in more than one "when" clause.
- Existential antecedents / preconditions - whether the domain's actions contain existential quantifiers in the preconditions or the effects.
- Object equality in antecedents / preconditions - whether the antecedents contained statements in which two parameters must be equal.
- Type hierarchy - whether the domain has nested types.
- implications - whether the actions contain implications in their preconditions.

In our experiments, we only used domains where our assumptions hold. In addition, SAM Learning of Conditional Effects (Conditional-SAM) does not support object equality checks in the antecedents thus, domains containing these conditions were not used in the experiments. Finally, type hierarchy is also not supported by Conditional-SAM since it can lead to ambiguities in predicate matching.

The domains in which all of the above holds are briefcase, cave-diving, citycar, elevators (miconic), maintenance, nurikabe and satellite. We did not use cave-diving in our experiments since both planners could not solve any of the problems and thus we could not create a trajectories dataset.

Safety Proof

Theorem 0.1. *The action model M' learned by Conditional-SAM is safe w.r.t the action model that generated the input trajectories $\mathcal{T}$.*

Proof. Conditional-SAM learns a superset of the original actions model's preconditions. Thus, for each action a and state s such that a is applicable according to M', it is guaranteed to be applicable according to M^*.

Given an action a and a state s in which a is applicable according to M', the resulting state $s' = a_{M'}(s)$, is equivalent to $s'^* = a^*_M(s)$, we prove this by contradiction. Assume that $s' \neq s'^*$. There can be two possibilities: (1) $\exists l \in s'$ such that $l \notin s'^*$, or (2) $\exists l \notin s'$ such that $l \in s'^*$. Since a is applicable in s according to M', it is also applicable according to M^*. Since Conditional-SAM only adds effects observed in the trajectories, according to rule 3, there cannot be a literal l such that (1) holds. If $l \in s'^*$ but $l \notin s'$, then Conditional-SAM did not observe l as a result of a and thus did not add it as an effect. According to line 15 one of ($l \vee$ *NotAnte*) hold in s. If $l \in s$, then $l \in s'$ according to M' (since a does not remove it), which contradicts (2). Similarly, if *NotAnte* $\subseteq s$, then the antecedent of l according to M^* is negated in s thus $l \notin s'^*$ which also contradicts (2). Thus Conditional-SAM learns a safe action model with respect to M^*. □

Space Complexity Proof

Theorem 0.2. *The space complexity of Conditional-SAM is* $O\left(|A||L|^{n+1}\left(\frac{e}{n}\right)^n\right)$.

For every action $a \in A$ and every literal $l \in L$, Conditional-SAM maintains the data structures *pre*(a), *PosAnte*(l,a) and *MustBeResult*(a). The size of *pre*(a) is at most $|L|$. The size of *MustBeResult*(a) is also at most $|L|$. The size of *PosAnte*(l,a) is observed when it is initialized, containing every conjunction of literals of size at most n, including the empty set (representing the antecedent *true*). Thus, the size of *PosAnte*(l,a) is at most $\sum_{i=0}^{n}\binom{|L|}{i} \leq \left(\frac{|L|\cdot e}{n}\right)^n$. We note that the space complexity of *BuildActionModel* is linear in the size of *PosAnte*. Consequentially, the space complexity is

$$|A||L| + |A||L| + |A||L|\sum_{i=0}^{n}\binom{|L|}{i} \in O\left(|A||L|^{n+1}\left(\frac{e}{n}\right)^n\right) \tag{1}$$

Recall that n is a fixed constant — the maximal number of literals in an antecedent.

Runtime Complexity Proof

Theorem 0.3. *The runtime complexity of Conditional-SAM is* $O\left(|A||L|^n\left(\frac{e}{n}\right)^n\right) + |\mathcal{T}||L|^{n+1}\left(\frac{e}{n}\right)^n\right))$.

Domain	Used in experiments	Planners can solve with the real model	Disjunctive antecedents (Assumption 4)	Existential preconditions / effects	Object equality in antecedents / precondition	Type hierarchy	Implications
Airport-adl	X	-	X	V	X	X	V
Briefcase	V	V	X	X	X	X	X
Caldera	X	-	X	X	V	V	X
Cavediving	X	X	X	X	X	X	X
Citycar	V	V	X	X	X	X	X
Elevators	V	V	X	X	X	X	X
maintenance	V	V	X	X	X	X	X
Nurikabe	V	V	X	X	X	X	X
Schedule	X	-	X	X	V	X	X
Satellite	V	V	X	X	X	X	X
Settlers	X	-	X	X	V	V	X
Spider	X	-	X	X	X	V	X

Table 1: Detailed information about the domains containing conditional effects available in the domains repository - https://github.com/AI-Planning/classical-domains

.

The initialization process requires the same runtime complexity as its space complexity, i.e., $O(|A||L|^{n+1}\left(\frac{e}{n}\right)^n)$. Then, Conditional-SAM iterates over all action triplets and applies the inductive rules in Def. 2. This requires $O(|\mathcal{T}||L|^{n+1}\left(\frac{e}{n}\right)^n)$, since as discussed above, the size of *PosAnte*(l,a) is at most $|L|^n\left(\frac{e}{n}\right)^n$.

Finally, in the *BuildActionModel* function, the runtime complexity is bounded by the most intensive computational part, which is the part that creates the restrictive conditions. The complexity of this part is linear in *PosAnte*(l,a). Thus the total runtime complexity of *BuildActionModel* is bounded by $O\left(|A||L|^n\left(\frac{e}{n}\right)^n\right)$ Thus, the total runtime complexity of the algorithm is $O\left(|A||L|^n\left(\frac{e}{n}\right)^n\right) + |\mathcal{T}||L|^{n+1}\left(\frac{e}{n}\right)^n\Big)$.

Analysis of Sample Complexity for Approximate Completeness

Proof. In view of Theorem 0.1, it suffices to show that for a pair $\langle P,\Pi\rangle$ drawn from $\mathcal{D}$, the preconditions of Π in M' are satisfied for each step of the execution of Π in the real action model M^*; indeed, the states obtained by M' and M^* are identical, so Π will then also solve P in M'.

Recall that Conditional-SAM passes the sets *pre*, *MustBeResult*, and *PosAnte* for each action a and, in the case of *PosAnte*, for each literal l to Algorithm 2. A literal l only appears in *pre*(a) for an action a if $\neg l$ has never been observed in the pre-state when action a was taken. Similarly, a clause $\neg c$ may appear as (a subclause of) some clause of the precondition *pre** of a in M' if c remains in the antecedents set *PosAnte*(e,a) of some candidate effect $e\in$ *MustBeResult*(a) for which more than one such candidate remains, or for which $e\notin$ *MustBeResult*(a) and c is in *PosAnte*(e,a). Note that if $\neg c$ is falsified in a state s (prohibiting a in s in M'), $s\subseteq c$. Hence, if the execution of Π in M^* would result in a being taken in s resulting in s', Conditional-SAM would remove c from *PosAnte*(e,a) for all $e\notin s'$, and c from *PosAnte*(e,a) if $c\nsubseteq s$ and $e\in s'\setminus s$.

We now claim that the probability that Conditional-SAM obtains a set of antecedents *PosAnte*(l,a) and set of preconditions *pre*(a) that prohibits the execution of Π with probability greater than ϵ is at most δ: a is only prohibited by *pre** in s if (1) $l\in s$ for some $\neg l\in pre(a)$; if (2) $s\subseteq c$ for some $c\in$ *PosAnte*(l,a) where $l\notin$ *MustBeResult*(a) and $\neg l\in s$; or, if (3) $l\in$ *MustBeResult*(a), $\neg l\in s$, $s\subseteq c$ for some $c\in$ *PosAnte*(l,a), and $s\nsubseteq c'$ for some (other) $c'\in$ *PosAnte*(l,a). When the execution of Π includes taking such an action a in such a prohibited state s, in the first case we see Conditional-SAM removes the falsified $\neg l$ from *pre*(a); for every effect e of a in s, any $c\nsubseteq s$ are removed from *PosAnte*(e,a) so cases (2) and (3) cannot occur; and for every literal $\tilde{e}$ that is not an effect of a in s, since $\tilde{e}\notin s'$, all $(c,\tilde{e})$ for $c\subseteq s$ are removed from *PosAnte*$(\tilde{e},a)$, so neither case (2) nor (3) can occur. Thus, we see that either at least one literal is deleted from *pre* or at least one c is deleted from some *PosAnte*(e,a) when such an (s,a,s') occurs in the trajectory, so that a is permitted by *pre*$^*(a)$ in s subsequently. Since literals are only deleted from *pre* and clauses are only deleted from *PosAnte*, Conditional-SAM then cannot return the eliminated collection of preconditions and antecedents sets.

Quantitatively, for any collection of preconditions and antecedents sets for which such a problem and plan would be obtained from $\mathcal{D}$ with probability greater than ϵ, Conditional-SAM can only return the corresponding collection with probability at most $(1-\epsilon)^m$ when it is given m examples drawn independently from $\mathcal{D}$. Observe that there are $3^{|F|}$ possible sets *pre* for each $a\in A$, and $2^{\sum_{k=0}^{n}2^k\binom{|F|}{k}}$ possible sets *PosAnte* for each l and a. Thus, there are at most

$$3^{|F||A|}2^{2|F||A|\sum_{k=0}^{n}2^k\binom{|F|}{k}}\leq e^{\ln(3)|F||A|+2\ln(2)|F||A|\left(\frac{2|F|e}{n}\right)^n}$$

possible collections of pre and *PosAnte*. Since $(1-\epsilon)^m\leq e^{-m\epsilon}$, taking a union bound over all possible collections of *pre* and *PosAnte* that prohibit the execution of the associated plan with probability at least ϵ, we find that for the given m, the total probability of Conditional-SAM obtaining such a collection of preconditions and antecedents sets is at most δ. Thus, with probability $1-\delta$, the action model indeed permits executing the plans associated with problems drawn from $\mathcal{D}$ with probability at least $1-\epsilon$ as needed.

Proof of Sample Complexity Lower Bound for Safe and Approximately Complete Algorithms

For any $p \geq 3|A|$, consider a domain in which there is a no-op action with no effects, and for each other action $a_i \in A$ there is a "goal" fluent f_i that is the effect of exactly one action, and this is the only effect. The domain includes an additional set of $(p - |A|)/2$ "flag" fluents, and $(p - |A|)/2$ "forbidden" fluents (so there are $|A| + 2(p - |A|)/2 = p$ fluents in total).

Now consider the following distribution $\mathcal{D}$ on problems and plans. The initial states have all goal fluents set to false, all but one (uniformly random) forbidden fluent true, and exactly n of the flag fluents (uniformly at random) true. With probability $1 - 4\epsilon$, the goal is empty. Otherwise, the goal includes a single goal fluent, chosen uniformly at random, that should be set to true. All other goal fluents, as well as the one forbidden fluent, must be set to false. The corresponding Π always consists of a plan with a single action; for the empty goal, the agent takes the no-op action, and otherwise the agent takes the action corresponding to the f_i goal fluent to be set true in the goal.

For any problem with a non-empty goal that we did not observe in the training set, the action model that is obtained from the true action model by adding the forbidden fluent as a conditional effect of the corresponding goal action with the flag fluents as the condition, is consistent with the training set. Indeed, either the action appears with a different set of flags so that one of the flag fluents in this condition is falsified (and the corresponding effect does not occur), a different forbidden fluent is false (so the relevant forbidden fluent is already true and the effect is not observed), or else the action differs from the one we need to achieve this goal, and then the effect is identical to the true action model. Therefore, no safe action model can permit taking the action needed to achieve the goal, and all other actions would reach a state in which some incorrect goal fluent is set to true and cannot be subsequently set to false.

Since the no-op goal only comprises $1 - 4\epsilon$ probability in the goal distribution, we need to observe at least a $3/4$ fraction of the possible goals for a safe action model to attain probability $1 - \epsilon$. But, there are $|A|$ goals, $(p - |A|)/3 \geq p/3$ forbidden fluents, and $\binom{(p-|A|)/3}{n} \geq \left(\frac{p}{3n}\right)^n$ sets of flags, and in expectation, a sample of size m only contains $4\epsilon m$ examples of these pairs of goals and flag settings. We, therefore, need $\Omega\left(\frac{1}{\epsilon}(|F|/3)^n|F||A|\right)$ examples; likewise, to even observe any of the nonempty goals with probability $1 - \delta$, we need $\Omega\left(\frac{1}{\epsilon}\log\frac{1}{\delta}\right)$ examples, giving the claimed bound.

□